Universidade de Aveiro
2012

Departamento de Biologia

Susana Manso Araújo

Identificação de fontes de poluição fecal na ilha da Berlenga

Tracking sources of fecal pollution in Berlenga island

Universidade de Aveiro
2012

Departamento de Biologia

Susana Manso Araújo

Identificação de fontes de poluição fecal na ilha da Berlenga

Tracking sources of fecal pollution in Berlenga Island

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Aplicada, realizada sob a orientação científica da Doutora Isabel da Silva Henriques, investigadora em pós-doutoramento do CESAM, Universidade de Aveiro e co-orientação da Doutora Anabela de Oliveira Pereira, investigadora em pós-doutoramento do CESAM, Universidade de Aveiro.

Apoio financeiro do FEDER no âmbito do programa COMPETE e da FCT- Fundação para a ciência e tecnologia através do projeto SEAGULL (PTDC/AAC-AMB/109155/2008 e FCOMP-01-0124-FEDER-008640).

o júri

presidente

Prof. Doutora Maria Adelaide de Pinho Almeida
professora auxiliar do Departamento de Biologia da Universidade de Aveiro

Doutora Daniela Rebelo de Figueiredo
investigadora em pós-doutoramento do CESAM, Universidade de Aveiro

Doutora Isabel da Silva Henriques
investigadora em pós-doutoramento do CESAM, Universidade de Aveiro

Doutora Anabela de Oliveira Pereira
investigadora em pós-doutoramento do CESAM, Universidade de Aveiro

agradecimentos

Quero expressar o meu especial reconhecimento a várias pessoas, que sem elas a realização deste trabalho não seria possível.

Ao Prof. Dr. António Correia, por me ter contagiado com o “bichinho” da microbiologia logo desde o meu primeiro ano de licenciatura e por me ter recebido no seu laboratório, sempre com o seu otimismo, força e motivação constante.

À Doutora Isabel Henriques por ter acreditado em mim, e por todo o apoio e incentivo constante ao longo destes anos. Agradeço o gosto que me incutiu pela investigação e por toda a preocupação demonstrada em momentos menos bons.

À Doutora Anabela Pereira agradeço toda a disponibilidade, sugestões e ajuda prestada que em muito contribuíram para a realização deste trabalho.

A todos do Instituto Politécnico de Leiria que se disponibilizaram ajudar-me, em especial ao Prof. Sérgio Leandro, Ricardo Nunes e Cátia Velez.

Aos melhores companheiros de laboratório com quem tenho o prazer de trabalhar, em especial, à Juliana, Nádia, Eliana, Dânia e Filipe, agradeço o ambiente de trabalho que me proporcionaram, e toda a amizade. Á Marta e à Isabel F. pelos momentos bem passados. Aos colegas de trabalho de outrora, João e Edna que me aturaram durante o meu primeiro ano de mestrado.
Em especial, à minha brasuca Ju, que me ensinou tudo o que havia para ensinar no meu primeiro ano no laboratório, e por continuar sempre ao meu lado, mesmo como amiga, e à Nádia, por toda a ajuda sincera que me prestou, por toda a paciência que teve comigo, e pela excelente ouvinte e conselheira que é.
A todos vocês agradeço o carinho, paciência e momentos especiais que passamos juntos.

Às companheiras de todo o sempre, Teté, Sónia e Rosa, agradeço-vos por todos os momentos caseiros fantásticos que partilhamos.

Às amigas para a vida que tenho a sorte e felicidade de ter, Silvana, Cláudia, Tati e Rosa agradeço pela amizade e apoio total, assim como todos os momentos de diversão. Sem vocês não teria ultrapassado um dos momentos mais difíceis da minha vida. Um muito obrigado Amigas!

Ao Bruno, que apesar das circunstâncias, me apoiou num período fundamental da minha vida, me ajudou a crescer, acreditou nas minhas capacidades e me ensinou a sorrir. Por tudo e por todo o sentimento que nos uniu, jamais esquecido, um muito obrigado!

À família, especialmente aos meus pais e à minha maninha, por todo o apoio incondicional que me têm proporcionado ao longo da vida, em particular, ultimamente.

palavras-chave

Poluição fecal, *Microbial source tracking*, Saúde pública*, Escherichia coli*, Ilha da Berlenga, Tipagem molecular

resumo

As águas marinhas costeiras são suscetíveis a contaminação fecal, tanto por fontes pontuais, como por fontes difusas, que podem ter contribuições de fontes individuais pertencentes a animais selvagens, animais domésticos e seres humanos. Os *inputs* de fontes difusas no ambiente são dispersos e esporádicos, o que torna a sua deteção difícil**.** A distinção entre a contaminação fecal de origem humana e não-humana tem vindo a tornar-se, nos últimos anos, um objetivo global crucial, uma vez que tem impacto na saúde humana e na economia local. Uma vez que a qualidade das águas superficiais é relevante para a saúde pública devido à sua ampla utilização, especialmente em atividades de lazer e consumo de marisco, a avaliação das fontes de poluição fecal primárias torna-se, assim, uma medida prioritária. Apesar da contaminação fecal por animais selvagens ser considerada de baixo risco para a saúde humana quando comparada com a poluição fecal de origem humana, as fezes de animais selvagens podem também transportar microrganismos patogénicos para humanos.

Nos últimos anos, um problema de contaminação fecal foi detetado na água da praia da Ilha da Berlenga. No sentido de esclarecer qual a origem desta contaminação surgiu este estudo, tendo como principal objetivo a determinação e identificação da(s) fonte(s) de poluição fecal responsáveis pela contaminação da água detetada na Ilha da Berlenga. Este objetivo foi alcançado utilizando a metodologia de "*Microbial Source Tracking"*, através de tipagem molecular (BOX-PCR) de isolados de *Escherichia coli* provenientes da água da praia, de fezes de gaivotas e de um efluente de origem humana e da análise dos dendrogramas resultantes. Para além disso, outros aspetos foram analisados, nomeadamente, a abundância relativa, a saturação de amostragem e índices de diversidade. Tendo em conta os dados resultantes do presente estudo, é possível concluir que: (i) as gaivotas podem ser consideradas o principal responsável pela poluição fecal da água praia; (ii) o método de amostragem e a estratégia da análise dos resultados obtidos podem ser considerados eficientes, para este tipo de ambiente e isolados; e (iii) o esforço de amostragem não foi suficiente para atingir toda a diversidade das populações de *E. coli* durante amostragem permitindo, no entanto, concluir quanto à principal fonte de contaminação fecal neste ambiente.

keywords

Fecal pollution, Microbial source tracking, Public health, *Escherichia coli*, Berlenga Island, Molecular typing

abstract

Coastal marine waters are often susceptible to fecal contamination from a range of point and nonpoint sources, with potential contributions from many individual sources belonging to wildlife, domesticated animals, and humans. These nonpoint source inputs into the environment are dispersed and sporadic, which makes their detection difficult. The distinction between human and non-human fecal contamination is becoming an important worldwide purpose, in light of the impact of fecal pollution on human health and economic affairs. Since quality of surface waters is relevant to public health due its wide use, particularly for recreational activities and seafood consumption, accurate assessment of primary sources of fecal pollution is clearly a priority measure. While fecal contamination from wildlife sources is often believed to present low human health risks compared to sewage, wildlife species are believed to carry human pathogens that may pose a health risk to humans as well.

In the last few years a problem of fecal contamination has been detected in the beach of the Berlenga Island. Thus, this study has emerged having as major aim the determination and identification of which sources of fecal pollution are the responsible for the water contamination detected in the Berlenga Island. This aim was achieved using a Microbial Source Tracking methodology through molecular typing (BOX-PCR) of *Escherichia coli* isolates from contaminated water, seagull feces and a human-derived effluent and analysis of the resulting clustering. In addition, relative abundance, sampling saturation and diversity indices were analyzed. Taking into account the data resulting from the present study, it is possible to conclude that: (i) the seagulls can be considered the main responsible for the fecal pollution of the beach water; (ii) the sampling method and the analysis methodology can be considered efficient to this type of environment and isolates; (iii) the sampling efforts were not enough to achieve all the diversity of the *E. coli* populations sampled allowing, however, the determination of the dominant source of fecal pollution in this environment.

INDEX

INDEX OF ABBREVIATIONS

A: adenine
AFLP: amplified fragment length polymorphism
ARA: antibiotic resistance analysis
BBR: Berlengas Biosphere Reserve
BOX-PCR: BOX elements – polymerase chain reaction
bp: base pairs
C: cytosine
CUP: carbon-source utilization profiling
cm: centimeter
DGGE: denaturing-gradient gel electrophoresis
dNTP's: deoxyribonucleotide triphosphates
EEC: European Economic Community
EC: European Council
EDTA: ethylenediaminetetraacetic acid
EU: European Union
ERIC: enterobacterial repetitive intergenic consensus – polymerase chain reaction
FAME: fatty acid methyl ester
FC: fecal coliforms
FS: fecal streptococci
FIO: fecal indicator organisms
FIB: fecal indicator bacteria
G: guanidine
H: Shannon diversity index
h: hour(s)
IUCN: International Union for Conservation of Nature and Natural Resources
J: equitability index
m: meter
MAV: maximum allowable value
$MgCl_2$: magnesium chloride
min.: minute(s)
mL: milliliter
mM: millimolar
MRV: maximum recommended value
MST: Microbial Source Tracking
NaCl: sodium chloride
$(NH_4)_2SO_4$: ammonium sulphate
no.: number
PCR: polymerase chain reaction
PFGE: pulse-field gel electrophoresis
RAPD: random amplified polymorphic DNA
rep-PCR: repetitive element PCR fingerprinting
REP-PCR: repetitive extragenic palindromic – polymerase chain reaction
RNA: ribonucleic acid

rRNA: ribosomal ribonucleic acid

T: thymine

TAE: tris-acetic-EDTA

T-RFLP: terminal restriction fragment length polymorphisms analysis

U: units of enzyme

UNESCO: United Nations Educational, Scientific and Cultural Organization

UPGMA: unweighted pair group method with arithmetic mean

UV: ultraviolet

V: volts

WHO: World Health Organization

WSPs: Water Safety Plans

w/v: weight by volume

WWTPs: wastewater treatment plants

ZPE: Zona de Proteção Especial

µg: microgram

µl: micro litter

µm: micrometer

µM: micro molar

°C: degrees Celsius

I. INTRODUCTION

1. BERLENGAS

The Berlengas archipelago is located on the Portuguese continental shelf, on the West side of Iberian Peninsula and Northwest of Cape Carvoeiro (Peniche). It distances of 5.7 miles approximately from the coastline, in a region characterized by two remarkable geomorphological accidents: the Cape Carvoeiro and the Nazaré Canyon. It is composed by three islands groups: Berlenga Grande Island and adjacent islets and reefs, Estelas and Farilhões Islands. These groups of islands extend to North-Northwest Berlenga, according to a length of slightly more than 4 miles (1,2).

Figure I.1. Location map of Berlengas Biosphere Reserve. (Based on Google™ earth data).

The archipelago has a land surface of approximately 104 ha, from which 78.8 ha corresponds to the Berlenga, the largest island, emerged area and additionally, 3.8 ha to the islets and reefs around it (1,2).

1.1 BERLENGAS BIOSPHERE RESERVE

Recently added to the World Network of Biosphere Reserves by the International Coordinating Council of UNESCO's Man and the Biosphere Program, the Berlengas archipelago and all the surrounding marine area was nominated as Berlengas Biosphere Reserve (BBR) (3). The actual BBR comprises the entire archipelago; the area emerged from the group of small islands and islets, as well as the adjacent marine area, with funds up to a maximum depth of 520 m. The total area of the BBR is 9541 ha, which is divided into 99 ha of land and 9442 ha of sea area (2).

Indeed, it is a fair recognition of the enormous potential and value of the natural heritage of the Berlengas archipelago since is only an additional title to that already owned.

Since September 1981 it is legally protected by the Decree-Law n. º 264/81 of September 3, being at this time classified as «Natural Reserve of Berlenga» by the Portuguese State, which has the 30 m bathymetric line as limit around the Berlenga and comprises all its islands, islets and sea area. Later in 1998 this area was reclassified being renamed as «Natural Reserve of Berlengas», consisting on the entire Berlengas archipelago and a Marine Reserve area, thus turning to be part of the national network of Protected Areas (Regulatory-Decree n. º 30/98, of December 23) (1,2).

In 1997 this area was integrated into the «Rede *Natura 2000*» under the Habitats Directive (92/43/EEC), and in 1999 was classified as «Zona de Proteção Especial (ZPE)» for the wild birds under the Wild Birds Directive (79/409/CEE), showing the value and importance of this area for biodiversity conservation at a European level. In addition to these statutes, the area is yet classified as a Biogenetic Reserve by the European Council (EC) (1,2).

The transversality of this high number of special designations that BBR owns has been widely recognized regionally and globally and proves the importance of this

archipelago as a single repository of genetic diversity, of species and habitats on the Western border of Europe, and its significance for the conservation of biological diversity. In addition, these statutes have at the same time a perspective of integrated management of various environmental, patrimony/cultural and socioeconomic aspects, since was clear at earlier 1981 the need to safeguard and enhance the patrimony, natural and cultural, of the land and sea, precisely where it assumes greater richness or where it is most vulnerable (1,2).

Biosphere Reserves have as the fundamental objective of promoting environmental sustainability through the creation of links between biodiversity conservation and economic development (2). Therefore there are specific aims taken as guidelines: (I) to promote the protection of the natural values of the archipelago and the surrounding marine area as the autochthonous flora and fauna and their habitats (1,4); (II) improve actions for management of human activities in that area, as part of a broader policy of marine conservation and sustainable use of the productive potential of the oceans, to preserve biodiversity and recover over-exploited or depleted resources (1,4); (III) and also, conjugate the management of this natural heritage with a perspective of sustainability ordering, controlling and improving its recreational, touristic and overfishing activities, enabling and promoting the sustainable development of economic activities (1,2,4); (IV) increase and share the scientific knowledge about marine and island communities (2).

1.2 UNIQUE CHARACTERISTICS AND BIODIVERSITY

The Berlengas archipelago hosts various peculiar forms with European and national relevance. From its insular nature to its geographical location and climate, complemented by a low and limited human occupancy motivated by the small size of the islands and land scarcity, contributed to the preservation and speciation of some of its unique features, like terrestrial and marine flora and fauna with several singularities, including some relevant ornithological aspects (1,2).

From the geological point of view the Berlengas archipelago consists in a complex of granitic and metamorphic rocks originated from the two supercontinents collision of Laurasia and Gondwana, during the Devonian and Carboniferous period.

As a result of this geodynamic process the islands and islets have a steep topography where is very common the formation of caves, and land and underwater cracks (2).

This set of coastal reefs is located in a temperate sea, under the influence of seasonal upwelling controlled by atmospheric circulation associated with the Azores Anticyclone, along of one of the most important submarine canyons in the international context, the Nazaré Canyon, and in the transition zone between the European and Mediterranean sub regions. This location is an important factor in the oceanographic dynamics of the region, mainly through the intensification of upwelling, renewal of nutrients to the surface and increased primary production, because contributes to the remarkable productivity and diversity of marine species and habitats (2,5).

The singular habitats arose possibly by the influence of two different climatic conditions: the Atlantic, in the northern cliffs and the Mediterranean climate in the southern cliffs (6). As habitats of particular significance at national and European level there are six habitats that are included in the European Union (EU) Habitats Directive. The most important ones comprise cliffs with vegetation of the Atlantic slopes, pioneer vegetation of *Salicornia* and other annual species from mud and sand zones, Mediterranean and thermo-Atlantic halophilous scrubs and halo-nitrophilous scrubs (2).

1.2.1 LAND AREA

In the land area of the Berlenga Island the presence of various species with high conservation value have been allowed by the maintenance and protection of a range of habitats which have a high ecological value too.

As result of a speciation process there are three endemic flora of great conservation value: *Armeria berlengensis*, *Herniaria lusitanica* subsp. *berlengianae* and *Pulicaria microcephala*. Apart from these, others species can be found, although not endemic, that have a restricted geographic distribution, being Iberian endemic or occurring only in the Iberian Peninsula and North Africa such as *Angelica pachycharpa*, *Calendula suffruticosa* subsp. *algarbiensis*, *Echium rosulatum*, *Linaria*

amethystea subsp. *multipunctata*, *Narcissus bulbocodium* subsp. *obesus*, *Silene latifolia* subsp. *mariziana*, *Silene scabriflora* and *Scrophularia sublyrata* (2).

Regarding to terrestrial fauna, specifically herpetological fauna, it is characterized by the presence of another endemic sub specie that has a high intrinsic value since it has particular characteristics derived from the insularity to which it is subjected, the lizard-of-Berlenga (*Podarcis carbonelli berlengensis*), abundant in Berlenga and Farilhões, and the ocellated-lizard (*Lacerta lepida*), only present in Berlenga represented as a residual population (1,2). The presence of terrestrial mammals is recorded only in the Berlenga island, and are example of that the wild rabbit (*Oryctolagus cunniculus*) and black rat (*Rattus rattus*), that were artificially introduced by Man (1,2).

With respect to avifauna, the Berlengas archipelago by its location in a region of high oceanic productivity constitute the limit south or north of nesting for three species of seabirds and the only nesting place in Europe of another specie. Moreover it has also an important role in the passage of migratory birds (e.g. *Hieraaetus pennatus* and *Luscinia svecica*), as the surrounding sea is an important feeding and concentration of sea birds area (2).

There are nesting records of seven species of seabirds in the archipelago, and all of these species have a conservation importance in the European context; they are the yellow-legged gull (*Larus cachinnans*), dark wing gull (*Larus fuscus*), tridactyl gull (*Rissa tridactyla*), Galheta or crested-cormorant (*Phalacrocorax aristotelis*), cagarra (*Calonectris diomedea*), airo (*Uria aalge*) and roquinho or alma-de-mestre (*Oceanodroma castro*); in addition, Berlenga island is the shelter of one of the largest western colonies of the yellow-legged gull (7).

1.2.2 MARINE AREA

The marine area of the BBR is very broad and characterized by high biological richness, also occurring here some species and habitats with high conservation importance in European and national context.

Given the location of the archipelago Berlengas on the continental shelf, ocean circulation in the vicinity of the islands is strongly influenced by wind and currents (2).

Many marine species perform their egg-laying on the continental shelf. These habitats provide great advantages for the development of larvae and juveniles, especially for the great abundance of food and high primary productivity, characteristic of these areas, refuge from predators and good conditions for rapid growth (2).

Thus, the oceanographic conditions of the marine area of the BBR give an abundant and diverse fish fauna, contributing to the presence of several species of marine mammals, particularly cetaceans, which include the presence of the bottlenose dolphin, *Tursiops corvineiro truncates* and the common dolphin *Delphinus delphis* (2).

The marine area is still characterized by rocky substrates and other substrates of biological origin (*Sabelaria* reefs), classified under the Habitats Directive, as well as benthic communities of plants and animals. Another important habitat contained in the Habitats Directive with high conservation value is the sea caves submerged or partially submerged. In the sea bottoms dominated by the occurrence of mobile sediments occur endobenthics, suspension and deposit feeders organisms (2).

The marine invertebrate fauna of the Berlengas archipelago is very diversified. The marine invertebrates more common in this marine area are the goose barnacles (*Pollicipes pollicipes*), octopus (*Octopus vulgaris*) and cuttlefish (*Sepia officinalis*), the limpet (*Patella intermediate*), anemones, shells, starfishes and nudibranchs, groups of species whom are highly sought after by scuba diving (8).

As mentioned by Rodrigues *et al.* (2008) there are referenced about seventy-six species of fish, and some, with high commercial interest. Among small pelagic are found the sardine (*Sardina pilchardus*), mackerel (*Scomber scombrus*), horse mackerels (*Scomber japonicus* and *Trachurus trachurus*), the conger (*Conger conger*), and some species of skates (*Raja* spp.) The most numerous family in terms of species is the Sparidae, with 11 sea breams species, as sargos (*Diplodus* spp.), pargos (*Pagrus* spp.) and goldfish (*Sparus aurata*), commercially important (8). Highlighting the conservation importance of the mero specie (*Epinephelus marginatus*) that it is

considered "In Jeopardy" by the International Union for Conservation of Nature and Natural Resources (IUCN), and much sought after, especially by practitioners for spear fishing (2).

1.3 CONSERVATION ISSUES

Beyond the recognition of the high value of the natural heritage of the Berlengas archipelago, this title demonstrates the recognition of the inherent problems of conservation and development, as well as the existence of specific operations to deal with such constrains.

In recent years, the local community and summer visitors have joined to actions that aim the nature conservation and natural heritage preservation, particularly through its strong commitment to the implementation of measures for more effective management in the archipelago, which undoubtedly in the future will revert to them. The island represents an *ex-libris* of tourism within local and regional area, exerting an highly attractiveness in the summer period, mainly in July and August; starting at the end of May the first temporary residents reach the island and end this demand in the middle of September (1,2,4). Taking into account the fragility of the ecosystem, tourism can be a problem to the conservation of the archipelago values. Because of that, despite of the number of people who may be at the same time on the island is legally regulated (270/90 of April 10), establishing the capacity of 350 visitors a day; unfortunately, there are periods of peak demand mainly during summer weekends, frequently leading to episodes of human overload. The high number of visitors in this season pressures on ecosystems and spawn general discomfort even for the visitors, causing pressure on the basic infrastructures of the island, including water supply, sanitation and waste production, and thus on the quality of life and local public health (2,4).

1.3.1 WATER SUPPLY AND SANITATION

With regard to infrastructure supply of fresh water, the Berlenga Island does not have its own resources, being currently only good storage conditions for

approximately 60.000 liters of fresh, without the warranty of wholesomeness for direct human consumption. Fresh water for drinking is bottled in land, being carried to the island by sea. Salt water is used for washing and to ensure the functioning of sanitary facilities, in all the houses of the fishermen's village, and restaurant. This water has to be first pumped from the sea to the supply deposits located in the highest elevation. Regarding the sanitation facilities they are rudimentary. Seawater is pumped into tanks and used in toilets, and later returned to the sea through ducts. Part of this salt water passes through a system of trituration waste and the washing water from the catering services are also released directly into the sea through the same pipeline system. There are no pits or other basic sanitation systems so that the release of these effluents directly to the sea may be reflected in the incidents of degradation of water quality and the occurrence of odors (2).

1.3.2 SEAGULLS

The avifauna is especially relevant for seabirds. The yellow-legged gull, *L. cachinnans* is by far the most abundant in the entire archipelago, particularly in the main island. This gull specie shows a clear expansion over time: estimations demonstrate that in 1939 the nesting population was about 2.000 individuals, having increased consistently for 32.000 birds in 1995, and then decreased until the present to just a bit more than 20.000 individuals. This reduction is the result of the implementation of various management measures, which consists on the eradication of mature individuals (measure taken in 1994-96) and destruction of their postures (since 1999). However, these management measures denote to be insufficient, since the issue remains. The current increase of the specie population on the island and consequent spread of the species to the mainland, throughout the Portuguese coast, is a problem of national importance (2).

2. MICROBIOLOGICAL WATER QUALITY

In the last years, there has been a human population increase and the spread-out of urbanization. The poor quality of water is a serious problem worldwide. About more than a billion people have no access to safe drinking water and millions die each year, suffering numerous waterborne infections after bathing in contaminated recreational waters (9). The natural aquatic ecosystems become microbiologically polluted mainly by point sources, through discharges of effluents from wastewater treatment plants (WWTPs), agricultural soil leaching as well as surface runoff, containing pathogenic organisms especially of fecal origin (10).

Moreover, recreational water activities which involve contact with water have grown in many countries worldwide. Also, nowadays, the ease of travel has altered the public use of water for recreational purposes, resulting in gradual deterioration of water quality (11). These recreational uses range from total-immersion sports, such as swimming, surfing and slalom canoeing, to non-contact sports, such as fishing, walking, bird-watching and picnicking (12,13).

Water is a natural resource that functions as an excellent carrier of numerous waterborne pathogens. Waterborne diseases arise, either by pathogenic bacteria, viruses and protozoa, or by chemical substances (14).

The fecal water contamination can cause a range of adverse health outcomes such as eye, ear, nose, and throat infections, through skin irritations, and finally it may be a reason of serious gastrointestinal illnesses or respiratory illness (11,13,15,16). In respect to contaminated bathing water it can cause serious and potentially fatal diseases (14). The number of waterborne outbreaks reported throughout the world has increased in recent years. This demonstrates the remaining of a significant cause of illness, although, the outbreaks are estimated based in detected cases which is likely to underestimate the problem (13,17,18). Furthermore, the number and type of pathogens in aquatic systems differs substantially depending of the incidence of disease among human and animal species and the seasonality of infection. Therefore, the numbers vary greatly between different parts of the world and times of the year (13,18).

In order to maintain water quality, numerous microbiological standards have been established. Throughout the world most of the countries have set up certain norms concerning water treatment and its final quality, on the basis of World Health Organization's (WHO) standards (11,13,15,19,20). Within EU there are four principal directives, enacted to manage the water policy within Member States, namely: the Urban Waste Water Treatment Directive (91/271/EEC) of 21 May 1991, the Drinking Water Directive (98/83/EC) of 3 November 1998, the New Bathing Water Directive (2006/7/EC) and Water Framework Directive (2000/60/EC) of 23 October 2000 (21–28). The microbiological monitoring of waters, intended to public use in EU countries, basically relies on the concept of microbiological indicators (21,26,28).

Monitoring of traditional fecal indicators, such as total or fecal coliforms, enterococci and *Escherichia coli* only indicates whether the body of water is impacted by fecal contamination. It does not provide any information on the source of such pollution, whereas this knowledge may help local authorities to restore water quality and reduce the risk of disease outbreaks (18). Therefore, Microbial Source Tracking (MST) approach has been spawned, which may not only assess water quality more accurately but also determine the source of contamination in water environment (17,18,29). The approach is based on the assumption that there are certain characteristics unique to the fecal microorganisms from specific hosts that may help to identify the source of fecal contamination (29,30). MST may discriminate the sources in broad fashion, like human vs. nonhuman sources; however group comparisons (humans vs. livestock vs. wildlife), species specific results (humans vs. cows vs. pigs etc.) as well as species individual hosts (cows from certain farm vs. other farms etc.) can be also performed (29). Currently, all of MST methods have several drawbacks, and there is no ideal MST technique that may be suggested as a standard for source tracking (29,31–33).

2.1 IMPORTANCE OF WATER QUALITY CONTROL

The maintenance of microbiological water quality is of special concern and imperative worldwide, as contamination of these systems can lead to high risks to human health and, as well as result in significant economic losses due to public health costs and closures of beaches and shellfish harvesting areas (10,14,34,35).

This include waters used for various purposes, namely water intended for drinking and used in food preparation, treated recreational waters as swimming pools, as well as untreated waters used for recreation like sea, river and lake water (13,36).

In natural aquatic systems the microbial quality of water is affected by various pathogens, including fecal bacteria, viruses and pathogenic protozoa (10,14,18). Recreational waters generally contain, besides the indigenous communities, a mixture of pathogenic and non-pathogenic exogenous microorganisms. These microbes may derive from several sources of contamination that can be point discharge of sewage effluents and industrial processes as well as non-point sources such as wastes of population water uses (defecation and/or shedding), animal husbandry (cattle, sheep, etc.), farming activities, leaching of soil and, the manure runoff (particularly in rural areas), and wildlife; in addition, recreational waters may also contain truly indigenous pathogenic microorganisms (10,12,13,37).

In the course of years, epidemiologists and microbiologists struggle with the fecal pollution problem, in order to protect public health from a number of outbreaks due to consumption of infected water and bathing in contaminated recreational waters. The problem is common to all nations regardless the economic status, but afflicts especially less economically developed countries (14,38,39).

Additionally, most of the waterborne illnesses remain undetected, and it is likely that, beyond the reported outbreaks, there is an unrecognized and underestimated problem (15). The WHO estimates that more than a billion people have no access to safe drinking water, and more than two million people, mainly children, die each year suffering infectious diseases associated with contaminated water (15,18,38). Moreover, global estimates suggest that specifically swimming and bathing activities within fecal polluted waters results in an excess of 175 million cases of infectious disease each year (39). Thus, basic hygiene-related diseases have a

significant impact on human health. The acute diarrhoeal disease is one of the most frequent causes of morbidity and mortality, causing alone 2.2 million of the 3.4 million water-related deaths per year, particularly children under five years old, being the second leading cause of death to this group (15,38).

Moreover, regarding to recreational waters, numerous epidemiological studies have been conducted worldwide in order to evaluate the association between recreational water quality and illness risk. These studies showed a positive correlation presenting higher rates of water related illnesses in swimmers compared with non-swimmers and assume that fecal indicator bacteria, in particular *E. coli*, can be used to predict gastrointestinal disorders, and in some cases, respiratory illnesses resulted from exposure to recreational waters (16,40).

The problem of fecal water contamination could be eliminated, or at least reduced, through the adoption of appropriate water quality practices, in particular, source protection and disinfection practices during potable water production and treatment of sewages (18). Nevertheless, especially in developing countries, the discharge of wastewater to the aquatic system still goes through partial or even no disinfection at all. Traditionally, the assessment of water quality is being performed through the analyses of fecal indicator organisms (FIO), also known as fecal indicator bacteria (FIB) to determine the microbiological quality of water, which are intentioned to indicate the presence of pathogens in water (18,38).

Numerous waterborne outbreaks have been reported throughout the world, not only viral, but also numerous bacterial and parasitic protozoa epidemics, involving fecal organisms such as *E. coli* O157:H7, *Campylobacter jejuni*, *Cryptosporidium*, *Giardia*, *Salmonella* and *Vibrio cholerae* (30,41). Although it should be stressed that the exposure does not always result in infection, nor does infection always lead to clinical illness, most of the illness contracted via recreational water are mild diseases, but a range of severities may also occur (11). The potential of microbial pathogens to cause ill in a considerable number of people increased the concern about water safety, and is well documented in countries from all levels of economic development. In 1993, Milwaukee (USA), was estimated about 400.000 individuals were affected by the outbreak of *Cryptosporidium*, a parasitic protozoan. Then, in 2000, in Walkerton, Ontario (Canada) there were reported over 2.600 cases which

resulted in six deaths, involving *E. coli* O157:H7. More recently, a serious outbreak turned out between October and November 2010 in South China. An acute gastroenteritis infection occurred due to the consumption of tap water contaminated with Norovirus contained in sewage (15,42). This demonstrates that not only developing countries have being afflicted with waterborne outbreaks.

The numerous outbreaks that are taking place every year and the demand for safe drinking water had generated and continues to generate, high social and epidemiological alarm. The WHO has been highly engaged with this issue. The preventive approach with important guidelines of universal application has been in development in order to monitor the quality of all water types (13,15). Even though several drawbacks recognition, classical FIO like coliforms, fecal (or thermotolerant) coliforms and *E. coli* have been useful along time, and are unquestionably the most commonly and successful used microbial parameters in drinking water quality assays, either by the ease of the assay or by the significant improvement it provided in the safety of drinking water all over the world (15).

2.2 PATHOGENS AND DISEASES

Water-based recreation and tourism can expose individuals to a variety of health hazards, including pathogenic microorganisms. Most of the microorganisms present in fresh and marine waters are not of concern to human health; still, some of them are responsible for some dangerous health outcomes. Typically, waters contaminated with human feces are regarded as a greater risk to human health, as they are more likely to contain human-specific enteric pathogens, including *Salmonella* spp., *Shigella* spp., hepatitis A virus, and Norwalk-group viruses. However animals can also serve as reservoirs for a variety of enteric pathogens (e.g., various serotypes of *Salmonella*, *E. coli*, and *Cryptosporidium* spp.) (17,34). In contrast, is known that most of the pathogens that are present in the human gastrointestinal flora do not colonize nonhuman species (43)

Water is not the natural habitat of pathogenic organisms, is solely a carrier of bacteria and viruses, which are introduced into aquatic systems directly from infected

humans or animals or indirectly through discharges of raw or insufficiently treated sewage and surface runoff of animal manure (13).

Depending on their characteristics, pathogens may cause asymptomatic or mild poisoning, and those transmitted through fecal-oral route, lodge in the alimentary canal, leading mostly to enteric infections, such as infectious diarrhoea (12,15). Though fecal-oral illness is not only caused by enteric bacteria, but also may result from virus or protozoa pathogenic (15). Epidemiological studies have revealed that contaminated water can be also a reason of serious gastrointestinal diseases, eye, ear, nose and throat infections, skin irritations, as well as respiratory system illnesses (11–13,15,16).

These morbidities can be orally transmitted through untreated or contaminated drinking water, but also due to bathing and other recreational activities, through inhalation, ingestion and/or skin penetration, in waters containing excrements (11,13,16). Some studies indicate higher risk health effects in swimmers comparing to non-swimmers and that FIO (in particular, *E. coli*) can be used to predict gastrointestinal and in some cases, respiratory illnesses from exposure to recreational waters (11,16). Additionally, children, elderly and immunocompromised individuals may be more predisposed to hazards as they are more susceptible to the pathogenic organisms that may occur in this environment (13). Furthermore, the illness risk from exposure to contaminated water may be significantly different between those associated with human sewage-impacted waters and by non-human sources (16).

More severe health outcomes may be associated with certain viruses, bacteria and protozoa, occurring among users of contaminated water, who are short-term visitors from regions with different rates of disease incidence (11,13).

2.2.1 PATHOGENIC BACTERIA AND PROTOZOA

Even though the majority of illnesses caused by waterborne bacteria are relatively mild, there are some bacterial and protozoa pathogens that may lead to acute health risks to humans or even leave sequelae (11,13). Both bacteria and protozoa may induce illnesses with a wide range of severity, once they occur as

parasites of animal guts: bacteria causing life-threatening diseases such as typhoid, cholera and leptospirosis, and on the other hand, protozoa may cause primary amoebic meningoencephalitis and dysentery (11,17).

They are, among others, bacteria from *Escherichia* spp., *Campylobacter* spp., *Mycobacterium* spp., *Salmonella* spp., *Leptospira interrogans*, *Shigella* spp., *Helicobacter pylori*, *Vibrio* spp.; and pathogenic protozoa such as *Entamoeba histolytica*, *Giardia lamblia* and *Cryptosporidium parvum* (11,44). From the most common bacteria reported within recreational waters, there are (11):

Escherichia spp. – although is not considered to be pathogenic once it colonizes human alimentary canals, this genus contains several strains like *E. coli* O157:H7 considered a worldwide concerning pathogen; has been associated with outbreaks, often reported from recreational waters.

Campylobacter spp. – in particular *C. jejuni* and *Campylobacter coli*, are the most common factor of bacterial gastroenteritis and chronic sequelae, linked to recreational waters.

Mycobacterium spp. – the species belonging to this genus are associated with a variety of diseases and can affect a variety of health conditions. *Mycobacterium ulverans* is generally pathogenic to healthy individuals, whilst *Mycobacterium avium* usually causes disease in immunocompromised individuals, such as skin and soft tissue infections, and respiratory related illnesses.

Vibrio spp. – there are evidence of *Vibrio vulnificus* infections association with recreational uses, when the user has a pre-existing open wound;

Regarding the protozoa pathogens:

Giardia – leads to giardiasis. The risk of death and the probability of developing sequelae from this infection are low, however in immunocompromised patients can be lasting as the acute illness can be prolonged and moderately severe.

Other pathogens, not so commonly associated with recreational waterborne diseases are *Legionella* spp., *Shigella* spp., *Salmonella* spp. bacteria, and *Cryptosporidium*, *Microsporidia*, *Naegeria fowleri* and *Schistosoma* spp. as protozoa.

2.2.2 VIRUSES

Viruses can cause serious diseases such as aseptic meningitis, encephalitis, poliomyelitis, hepatitis and myocarditis. These enteric viruses are present at high densities in human waste and therefore they might enter the water bodies through discharge of sewage contaminated water (11,44).

Some of the most important fecal viral pathogens are noroviruses, enteroviruses, adenoviruses, rotaviruses, and hepatitis A and E viruses (11,15,44–46). The echovirus infections in recreational water may occur through fecal contamination, and there may be a lot of infections cases unreported. The transmission of adenovirus may occur through swimming in fecal-polluted recreational waters, resulting in conjunctivitis, pharyngitis, pneumonia, acute and chronic appendicitis, bronchiolitis, acute respiratory disease, and gastroenteritis (11).

Overall, viruses are more resistant to environmental conditions than bacterial indicators, which in part explain the frequent lack of correlation between currently used indicators and the occurrence of enteric viruses (11,15,44,46).

The review of the most prevalent waterborne pathogens with related diseases and its source are presented in the Table I.1 (11,44,46,47).

Table I.1. Examples of common waterborne pathogens, some related diseases and their sources. Adapted from (11,44,46,47).

Pathogen	Disease	Reservoir/Source
Bacteria		
Campylobacter spp.	Gastroenteritis	Human and animal feces
H. pylori	Gastroenteritis, ulcers, anaemia, gastric cancer	Human and animal feces
Salmonella spp.	Gastroenteritis, typhoid fever, salmonellosis	Human and animal feces
Shigella spp.	Bacillary dysentery	Human feces
V. cholera	Cholera	Human feces
E. coli	Gastroenteritis	Human and animal feces
Yersinia enterocolitica	Gastroenteritis	Human and animal feces
Leptospira spp.	Leptospirosis	Animal and human urine
Virus		
Enteroviruses	Poliomyelitis, aseptic meningitis, hemorrhagic conjunctivitis, myocarditis, encephalitis, herpangina	Human feces
Rotaviruses	Gastroenteritis	Human feces
Adenoviruses	Upper respiratory and gastrointestinal illness	Human feces
Hepatitis A and E virus	Hepatitis, miscarriage and death	Human feces
Norovirus	Gastroenteritis	Human feces and water
Protozoa		
Acanthamoeba castellanii	Amoebic meningoencephalitis	Human feces (not strict)
Balantidium coli	Balantidosis (dysentery)	Human and animal feces
Cryptosporidium homonis, *C. parvum*	Cryptosporidiosis (gastroenteritis)	Water, human and other mammal feces
E. histolytica	Amoebic dysentery	Human and animal feces
G. lamblia	Giardiasis (gastroenteritis)	Water and animal feces

2.3 EUROPEAN UNION'S STANDARDS

The EU has decreed four principal directives in order to manage the water policy within member countries, namely:

- The Urban Waste Water Treatment Directive (91/271/EEC) of 21 May 1991 concerning the discharges of municipal and industrial wastewaters (24);

- The Drinking Water Directive (98/83/EC) of 3 November 1998 relating to potable water quality (26);

- The New Bathing Water Directive (2006/7/EC) of 24 March 2006 concerning the healthiness of bathing waters (21,27);

- Water Framework Directive (2000/60/EC) of 23 October 2000 regarding water resources management (28).

The Urban Waste Water Treatment Directive (91/271/EEC) regards the collection, treatment and discharge of urban wastewater from both domestic and certain industrial sectors. It has as main objective the protection of the environment from any adverse effects caused by discharge of urban wastewaters. According to the Directive, Member States shall ensure that all population agglomerations may provide the appropriate collection of urban wastewater, and before discharge be subjected to proper treatment (24).

The Drinking Water Directive (98/83/EC) concerns the quality of all water intended for human consumption, apart from mineral and medicinal waters and small water supplies (<50 persons). The Directive's objective is to assure that Member States provide their customers with clean and wholesome water, free from any microorganisms, parasites and any other substances that may constitute potential risk for human health. The distribution network water has to meet the minimum microbiological requirements established by the Directive, while the Member States shall take all measures necessary to guarantee the beneficial effects and purity of water and avoid risks for public health through regular monitoring of the water quality using the methods specified in the Directive. The monitoring consists in the control of potentially damaging substances in drinking water, through a set of established relevant parametric values that must be subject to check monitoring in samples representative of the quality of the water consumed throughout the year

(26). The European Union has initiated in 2003 a major revision of the currently Drinking Water Directive (98/83/EC) to decide which modifications have to be included in the new and updated Directive in order to increase the quality of drinking water, and protect public health. The new approach, named "Water Safety Plans" (WSPs) takes into account not only the quality of the finished water, but also wants to encompass the water quality from the source to the final tap water at the consumer's home, once several entities and agencies have encountered some cases in which the presence of indicators was detected at the time of serving the water to the consumers (48).

The New Bathing Water Directive (2006/7/EC) was adopted to preserve, protect and improve the environment quality and to protect public health from the risk of accidental and chronic diseases caused by human pathogens present in contaminated water (27). The procedure for monitoring of recreational waters established by this revised Directive gives more reliable results, as it requests for stringent water quality standards and reduces the list of nineteen bacterial indicators that need to be monitored to just two microbial indicators of fecal contamination, namely, *E. coli* and intestinal enterococci, replacing the assessment under the first European Bathing Water Directive from 1975 (76/160/EEC), which was based on percentage compliance counts of fecal index organisms (13,21,49,50). Additionally, it puts a stronger emphasis on beach management and public information. In 2015 the current Directive will be repealed to the revised Bathing Water Directive (2006/7/EC) come into force (27,50).

The EU Water Framework Directive (2000/60/EC) was established in 2000 in response to the increasing pollution and increasing demand for clean rivers, lakes and beaches throughout Member States. The aim of this Directive is to maintain and improve all the aquatic environments and provide the framework for protection and improvement of a quality of all types of waters (among others, rivers, lakes, estuaries, coastal waters, groundwater), and the sustainable use of water in the European Community. It sets a clear objective of by 2015 all European waters may be in "good status" (28,51).

All Member States of the EU are legally obliged to comply with the standards laid down in these Directives, including Portugal. They have to assure the minimum

quality standards of the various aquatic bodies, beginning through drinking water wholesome, through recreational water healthiness and finally setting standards for urban wastewater treatment in each country. For this, each country has to monitor the quality of their water using standard methods, to reduce pollution of water bodies and protect against deterioration (21,22,24,26,28).

2.3.1 PORTUGUESE STANDARDS

In Portugal the water quality is established by the European Standards, as well as by internal legal Decrees, namely the Decree-Law 306/2007 of 27 August 2007 concerning the water intended for human consumption; the Decree-Law 236/98 of 1 August 1998 regarding the waters to produce drinking water, aquaculture and irrigation water; and the Decree-Law 135/2009 of 3 June 2009 related to bathing waters (52).

The Decree-Law 306/2007 concerns the water intended for human consumption and appeared to transpose into the Portuguese legal system the European Drinking Water Directive (98/83/EC). This Decree-Law has the aim of ensure universal availability of clean water, and greater balance in its composition in order to protect the human health from the possible harmful effects that may result from contaminated water supply. In addition, this Decree defines the methods for monitoring the water quality, establishing a routine inspection which defines the location of sampling points and the minimum sampling frequency (52).

The main objective of the Decree-Law 236/98 is to establish standards, criteria and quality objectives in order to protect the aquatic environment and improve water quality in terms of its main uses. The water for human consumption here considered is of two types: groundwater and fresh surface water intended for the production of drinking water (52).

The bathing waters Decree-Law (135/2009) was created in 2009 in order to transpose into the Portuguese national law the New Bathing Water Directive (2006/7/EC) of the European Parliament and Council, and complementing the Water Law, approved on 29 December (Law no. 58/2005). The main purpose of this order is to establish the legal framework for the identification, management, monitoring and

classification of bathing water quality and the supply of information about them to the public. Bathing waters are defined as surface waters, whether inland, coastal or of transition, as defined in the Water Law (Law no. 58/2005) that may expect a large number of bathers and where bathing has not been banned or discouraged permanently. Monitoring frequency should be conducted as stated in the Decree-Law 135/2009, as well as the microbiological parametric values to be followed in assessing the quality of inland bathing waters. The monitoring parameters to evaluate the microbiological assessment of bathing water quality are the same for inland waters and for coastal and transitional waters; however, the values of such parameters differ between the two types of water (Table I.2) (52).

Table I.2. Microbiological parametric values to be followed in assessing the quality of inland, coastal and transitional bathing waters. Adapted from (52).

Water Type	Parameter	Unity	Quality		
			Excellent	Good (MRV)	Acceptable (MAV)
Inland	Fecal Enterococcus	Number/100 mL	200	400	330
	E. coli	Number/100 mL	500	1000	900
Coastal/ Transition	Fecal Enterococcus	Number/100 mL	100	200	185
	E. coli	**Number/100 mL**	**250**	**500**	**500**

MRV, maximum recommended value that should not be exceeded

MAV, maximum allowable value that must be respected or not exceeded

Therefore, range of values between MR and MA values means that will be no significant risks to bathers health (Decree-Law 236/98) (52).

3. MICROBIAL SOURCE TRACKING

Traditionally, the evaluation of water quality and health risk is made by cultivation and enumeration of FIO, such as total coliforms, fecal coliforms, *E. coli* and enterococci (29,30,53). These microorganisms are normally present in the intestinal tract and feces of warm-blooded mammals, including wildlife, livestock, and humans, thus, the indicator bacteria themselves are usually not pathogenic (18,30); additionally they are not able to survive and multiply in this environment. Their presence in water solely indicates the potential presence of enteric pathogens within the contaminated water, and therefore they are used in monitoring as they are much easier and less costly to detect and enumerate than the pathogens themselves (29,30). An ideal indicator would be non-pathogenic, rapidly detected, easily enumerated, and have similar survival characteristics to the pathogens it indicates (29,30).

In the past, attempts to classify fecal sources based on FIO focused on discriminating contamination sources in a broad fashion (i.e., human vs. nonhuman categories) based on the fecal coliforms/fecal streptococci (FC-FS) ratios (29,30). Although these classical approaches failed to accurately differentiate the source of fecal pollution between human and animal sources of pollution (once the FC-FS ratios are not consistently valid for different animals), it is recognized that the knowledge about the fecal source can help local communities to restore water quality and reduce the risk of outbreaks (29,30).

Microbial Source Tracking (MST) is a rapidly emerging area from a growing need to determine the source(s) of fecal contamination impacting a water system. MST method may not only assess water quality more accurately, but also determine the source of fecal pollution, given the appropriate method and fecal source identifier (17,29,30). MST is based on the assumption that there are unique strains of microorganisms adapted to their specific hosts, and with the help of these differences, the source of fecal microbial contamination can be identified (29,30).

Besides all of the above-mentioned qualities to an ideal indicator, an ideal MST microorganism must have as well discriminatory power between hosts (30).

Currently, the possible source discriminations are: broad fashion discrimination (human vs. nonhuman sources), species specific results (humans vs. cows vs. pigs etc.), host group comparisons (humans vs. livestock vs. wildlife), and specific individual hosts (cows from a certain farm vs. other farms vs. other livestock on farms vs. human etc.) (29).

There are various methods that can be used to seek for the origin of contamination in water. Some of these methods intend the discrimination between human and non-human sources of fecal contamination, and some others are designed to differentiate between fecal contaminations originated from more than two animal species (31). The currently MST methods employ genotypic or phenotypic characterization of microorganisms from water bodies. Genotypic analyses rely on certain aspects of organism DNA sequence, whereas phenotypic assays measures specific feature that is expressed (29,30). The methods are further divided into library-dependent and library-independent, from which some require cultivation of target organisms and the others are culture-independent (29,30). Despite the fact that comparison studies have demonstrated that no single method has been found undoubtedly superior to another, MST methods and technology are still being developed (29–33).

3.1 LIBRARY-DEPENDENT/CULTURE-DEPENDENT METHODS

The library-dependent methods are based on a host-origin collection of isolates from known fecal sources, so called reference library. The microorganisms isolated from unknown sources are analyzed in order to provide a set of "fingerprint" patterns that are further compared with the isolate profiles of the reference library, classifying indicator organisms of unknown origin by source category (29). The effectiveness of all library-based methods strictly depends on the size and representativeness of the library in a known-source, although the size factor needs to be addressed. Moreover, it should be stable over time so that there is no need to continually create new libraries. The library-dependent methods include both phenotypic and genotypic tests and are culture-based (17,29,30).

Phenotypic characterization measures a trait and includes: antibiotic resistance analysis (ARA), carbon-source utilization profiling (CUP), Fatty acid methyl ester (FAME) profiling (17,18).

Genotypic methods rely on molecular typing or fingerprinting techniques that are used to differentiate specific microorganisms. This approach includes: repetitive element PCR fingerprinting (rep-PCR), random amplified polymorphic DNA (RAPD) analysis, amplified fragment length polymorphism (AFLP) analysis, pulse-field gel electrophoresis (PFGE) and ribotyping (18,29).

3.2 LIBRARY-INDEPENDENT/CULTURE-DEPENDENT METHODS

The library-independent culture-based approach is relatively simpler than the library-dependent's methods, since it is not required a library. It relies on presence or absence of a target organism or gene in the sample. When the target for MST analysis is in low number, it is necessary to primarily enrich the sample or obtain isolates. This methodology comprise F+RNA coliphage typing and gene specific PCR (17).

3.3 LIBRARY-INDEPENDENT/CULTURE-INDEPENDENT METHODS

Cultivation-independent methods are principally based on nucleic acid analyses. They employ a genetic marker from DNA extracted from water sample, without any culturing procedure. The great advantage of this approach is the quickness of the process and the no need of a library, as the markers are universal in most of the cases. On the other hand, currently the markers are limited to host species beyond humans and a few important domestic animal species (17,18).

Genotypic characterization can be done by total community analysis, through 16S rRNA gene clone libraries identification, community fingerprinting, by denaturing-gradient gel electrophoresis (DGGE) or terminal restriction fragment length polymorphisms analysis (T-RFLP) and, target specific PCR-based method (host-specific markers and virus specific markers) (17,18).

The Table I.3 presents a review of the methods applied in MST, and their advantages and disadvantages (29,30,34,45).

Table I.3. Advantages and disadvantages of current methods used for MST methodologies. Adapted from (29,30,34,45).

Method	Advantages	Disadvantages
ARA	- Rapid and easy to perform - Requires limited training - High discrimination power	- Reference library required - Requires cultivation of target organism - Libraries geographically/temporally specific -Variations in methodology in different studies
CUP	- Rapid and easy to perform - Requires limited training	- Reference library required - Requires cultivation of target organism - Libraries geographically/temporally specific -Variations in methodology in different studies - Results often inconsistent
rep-PCR	- Highly reproducible - Rapid and easy to perform - Requires limited training - High discrimination power	- Reference library required - Requires cultivation of target organism - Libraries geographically/temporally specific - Variability increases as library increases
RAPD	- Rapid and easy to perform - High discrimination power	- Reference library required - Requires cultivation of target organism - Libraries geographically/temporally specific - Not been used extensively for source tracking
AFLP	- Highly reproducible - High discrimination power - Can be automated	- Reference library required - Requires cultivation of target organism - Requires specialized training of personnel - Labor-intensive - Expensive equipment required - Libraries geographically/temporally specific -Variations in methodology in different studies
PFGE	- Highly reproducible - High discrimination power	- Reference library required - Requires cultivation of target organism - Requires specialized training of personnel - Labor-intensive - Libraries geographically/temporally specific
Ribotyping	- Highly reproducible	- Reference library required

	- High discrimination power - Can be automated	- Requires cultivation of target organism - Requires specialized training of personnel - Labor-intensive (if not automated) - Libraries geographically/temporally specific -Variations in methodology in different studies
F+RNA coliphage	- Discrimination between human from animals - Subtypes are stable - Easy to perform - Reference library not required	- Requires cultivation of coliphages - Subtypes do not exhibit absolute specificity - Low in numbers in some environments
Gene specific PCR	- Can be adapted to quantify gene copy number - Virulence genes may be targeted, providing direct evidence of harmful organisms are present - Reference library not required	- Require enrichment of target organism - Sufficient quantity of target genes may not be available requiring enrichment or large quantity of sample - Requires training of personnel - Primers currently not available for all relevant hosts
Community fingerprinting	- No cultivation required - Rapid and easy to perform - Relatively inexpensive - Reference library not required - Host specific	- Portion of community that can be linked to host specificity may be very small compared to indigenous microbial community - Has not been widely used for MST
Host-specific markers	- No cultivation required - Rapid and easy to perform - Reference library not required - Indicator of recent pollution	- Little is known about survival and distribution in water systems - Primers currently not available for all relevant hosts - Control measures required to avoid cross-contamination
Virus specific markers	- Host specific - Easy to perform - Reference library not required	- Low in numbers, requires large sample size - Not always present even when humans present

The selection of most appropriate method for tracking of fecal pollution source depends on several factors, such as: complexity of aquatic system, level of contamination, bacterial strains applied for tracking, character of investigation (human/non-human or differentiation between animal species), availability of resources (funds, time constraints, personnel with technical knowledge, equipment) and time (sample processing and data analysis) (17,29). At present, there is no ideal recommendation of which MST approach and method to apply to all fecal pollution source tracking situations. More research needs to be addressed to minimize the issues related to the available techniques. The use of a toolbox of methodologies rather than a single approach is also being studied (17,18,29,30).

This study employed the repetitive element sequence-based polymerase chain reaction (rep-PCR) that is a well-established technology in diversity analysis of very closely related species, especially in differentiation of bacteria strains within one specie (54).

This technique relies on the differentiation between different pollution sources using repetitive intergenic DNA sequences elements that many bacterial species have distributed in multiple copies throughout their genomes (54). These repetitive elements are thought to be highly evolutionarily conserved because are essential protein-DNA interaction sites or because these sequences may propagate themselves as “selfish” DNA by gene conversion (55). This method is one of the commonest methods used to identify sources of fecal contamination of water system (56). In order to produce DNA fragments of various sizes, the DNA flanking the repetitive extragenic elements is amplified using polymerase chain reaction (PCR) and rep-specific primers (57). Amplification of the distinct genomic regions located between these repetitive elements results in a distinctive strain pattern (58). The resulting amplicons are separated by electrophoresis and the strain-specific DNA fingerprints can be analyzed through recognition of patterns and their comparison with the library to determine the genetic relatedness (57,59). Bacteria, which have identical fingerprints, are considered as being the same strain, or clonal, while those possessing similar patterns are considered as being genetically related (54,60).

4. MICROBIAL INDICATORS OF FECAL POLLUTION

Indicator microorganisms are used to predict the presence of pathogenic microorganisms and minimize the potential risk associated. The use of FIO is advantageous in the way that they bypass the need of assay for every pathogen that may be present in water. An ideal indicator may be: an inhabitant of the gastrointestinal tracts of warm-blooded animals; non-pathogenic to humans; present in greater numbers than the pathogen, and should be absent in uncontaminated samples; easy, rapid, and inexpensive to detect and enumerate than the pathogens themselves; have survival characteristics similar to the pathogens of concern, but not multiply in the environment though; be at least equally resistant as the pathogen to the environmental factors and to disinfection in water and wastewater treatment plants; be distributed randomly in the mass of water, and can be strongly associated with the presence of pathogenic microorganisms (34,37,52).

Consequently, the identification of the sources of these enteric bacteria has been the focus of most MST studies thus far, primarily because they are the basis of microbial water quality criteria, secondly because they are considered common inhabitants of most endothermic animal guts, and lastly due to their relatively ease to culture. Additionally, selective media are available for their isolation, which minimizes the number of false positives that need to be further characterized (17,29).

E. coli has been used for long as an indicator of fecal pollution. It has good characteristics of a fecal indicator, such as not normally being pathogenic to humans, and is present at concentrations much higher than the pathogens it predicts. However, recent studies have provided evidence suggesting that *E. coli* may not be a reliable indicator in tropical and subtropical environments due to its ability to multiply in the environment (15,29,34). Nevertheless, total and fecal coliforms, such as *E. coli* and *Enterococcus* spp., have been used extensively for many years, as indicators for determining the quality of all types of water bodies in the vast majority of MST studies, and continue to this date, being *E. coli* the most commonly found within literature and for that considered the best indicator for fecal contamination (17,34,61).

4.1 *ESCHERICHIA COLI*

E. coli is a fecal coliform that has been extensively used as an indicator because of being an inhabitant of the intestinal tracts and excreted by all warm-blooded animals; although it presence in water does not indicate the source of fecal pollution (43,57). In fresh feces it can achieve concentrations of 10^9 per gram (43,52,61). Most of the *E. coli* found in the human gut are harmless but there are about five pathogenic groups of *E. coli* that may cause disease in humans (11).

It is possible to find *E. coli* in sewage, effluents after treatment, and in general in all types of waters and soils recently subjected to fecal contamination, whether by humans, wild animals, or agricultural activity. As a result, pathogenic organisms human-infectious may be transmitted by wild animals, including birds, even in the remotest regions. Thus, the presence of *E. coli* must not be ignored in any places fecal contaminated because its existence is a sign that the water has been contaminated and a potential treatment has been ineffective. The detection of *E. coli* in water after the water body been submitted to treatment is of the same significance as any other coliform organism, but its absence does not mean that pathogens have been eliminated, as like the other coliform indicators, it is more sensitive to disinfection practices than many associated pathogens (in particular viruses). However, it is widely used as an indicator of treatment effectiveness (15).

In addition to the traditional microbiological indicators, there are many other indicator microorganisms, pathogens, or chemical markers that can be used in a methodological approach of MST. Moreover, independently of which fecal identifier is selected, it can always be complemented by other method.

5. AIMS OF THE WORK

The distinction between human and non-human fecal contamination is becoming an important worldwide purpose, in light of the impact of fecal pollution on human health and economic affairs. Since quality of surface water is relevant to public health due its wide use, particularly for recreational activities and seafood production and consumption, accurate assessment of primary sources of fecal pollution is clearly a priority measure.

While fecal contamination from wildlife sources is often believed to present low human health risks compared to sewage, wildlife species can carry human pathogens that may pose a health risk to humans as well.

In the last few years a problem of fecal contamination has been detected in the surrounding sea water of the Berlenga Island, and even featured in the social communication, leading to the shore closure and associated revenue losses in the local economy, since it depends on the touristic activities.

In this order of ideas, our research group leads an FCT funded project "Identification of non-point sources of fecal pollution in a natural environment: contributing data for risk assessment" (SEAGULL), reference PTDC/AAC-AMB/109155/2008, that aims to identify the source of the fecal pollution detected and to assess the human health risk.

Thus, this study has emerged for this purpose, having as major aim the determination and identification of which source of fecal pollution is responsible for the water contamination detected in the Berlenga Island beach.

As a secondary output of the work, beyond the main aim, this study will enable the establishment of an *E. coli* collection from three different sources: beach water, seagull feces and human effluent. This collection will be used for future studies regarding the assessment of risk of the fecal pollution detected in the Berlenga's water.

II. EXPERIMENTAL PROCEDURE

1. WATER QUALITY INDICATORS

Samples of water from the beach were collected during summer season from May to September of 2011. The sampling of water was done every week. Then the samples were sent to Quimiteste laboratory (Palmela, Portugal) in order to perform *E. coli* and Enterococcus counting.

The water harvesting consisted in collecting 500 mL into a sterile vial at about 30 cm from the surface, in a location where the water column was at least 1 m deep, for both high tide and low tide (52). The vials were then stored at 4°C, until the return to the laboratory.

2. SAMPLE COLLECTION

The sampling was performed every two weeks from May to September of 2011 (unless setbacks). Samples were collected from three locations of different sources, as can be seen in figure II.1. It was collected water of the Berlenga beach «Carreiro do Mosteiro», seagull feces scattered in the beach and/or of the surrounding rocks, and effluent derived from the island sanitary infrastructures. All the samples collected were then properly stored at 4°C in the fridge until the return to the mainland (ESTM laboratory, Peniche), within refrigerating coffers.

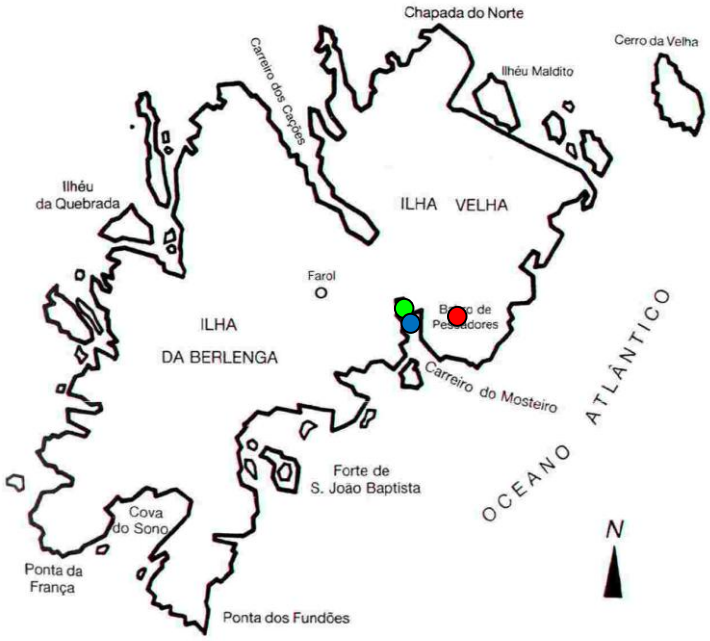


Figure II.2. Location of the sampling sites. Blue spot corresponds to water; green spot corresponds to feces; and red to effluent. Adapted from (87).

2.1 WATER

A volume of 2 L of water samples was collected at high tide with a sterile vial at about 30 cm from the surface, in a location where the water column is at least 1 m deep. The vials were then stored at 4°C till used. Sampling events were timed to coincide or to be near high tide.

2.2 FECES

In each sampling date 5 composite samples were collected composed by 5 to 10 individual fecal samples. Samples were collected with the help of a spatula into a sterile tube and after stored at 4°C till the return to the laboratory. The samples were collected as early as possible in the morning before tourists reach to the beach.

2.3 EFFLUENT

Effluent was obtained from an opening in the pipes of discharge to the sea in the morning or afternoon discharge moments. It was collected approximately 250 mL into a sterile vial and stored at 4°C till the return to the laboratory.

3. SAMPLE PROCESSING

The first moment of sampling (campaign I) served to optimize the dilution factors for fecal and effluent samples and selection of sample volumes (in case of water and effluent type samples). All samples were analyzed by using a membrane filtration method (62).

3.1 WATER

The samples of water were subjected to filtration in three volumes: 10, 30 and 50 mL. Three replicates for each volume were filtered through a 0.45 µm membrane filter (Millipore, USA) that retains the bacteria, under a vacuum system, prior to filter placing into the surface of Chromocult® Coliform Agar (Merck, Germany) plates and it's incubation at 37°C during 18 to 24 h.

3.2 FECES AND EFFLUENT

The 5 fecal composite samples were subjected to homogenization. Further 1 g of these homogenized samples was used in the preparation of a solution with 100 mL of saline solution to use as solution base for the decimal dilutions. Effluent samples were also subjected to serial dilutions in saline solution (0.9% w/v of NaCl).The decimal dilutions of range 10^{-2} to 10^{-8} were passed through a 0.45 µm membrane filter, under a vacuum system. Membranes were then put into plates of Chromocult® Coliform Agar (Merck, Germany), and further incubated at 37°C during 18 to 24 h. All the plates were properly transported to the laboratory within a cooler bag.

4. *ESCHERICHIA COLI* ISOLATION AND PURIFICATION

After overnight incubation at 37°C, one plate of the three replicates of each condition was chosen to further proceed to *E. coli* strains isolation of each source.

For each time of sampling, all the colonies with dark-blue to violet aspect were selected, resulting from the selective and differential chromogenic nature of the culture medium CCA for detection of total coliforms and *E. coli*. Each presumptive *E. coli* colony was streaked onto the surface of CCA plates and incubated overnight for 18 to 24 h at 37°C. Being this step repeated as many times as necessary until pure cultures were obtained. Once every single colony was purified in the selective medium, each *E. coli* strain was streaked on Tryptic Soy Broth agar plates, for maintenance of the collection.

4.1 *ESCHERICHIA COLI* CONFIRMATION

Several colonies, about 1% of the final number of isolates, were checked with the Kovac's indole reagent test, and others selective/differential media like MacConkey and mFC agar, to confirm the identity of *E. coli*. These media showed that we were in the presence of *E. coli* strains intended, as in the first one was obtained pink colonies, and with the second one, blue colored colonies were checked. As for the Kovac's indole test turned the colonies into cherry-red color, confirming the presence of *E. coli*. Moreover, for 4 colonies the confirmation of the identity of the presumptive *E. coli* isolates was achieved through sequencing the total 16S rRNA gene and comparison with sequences in public databases. For this, one colony of each isolate was picked and resuspended in 20 µl ultra-pure distilled water and heated at 100°C. Thus, 1 µl of the suspension was used as DNA template to 16S rRNA gene amplification through PCR. The 16S rRNA reactions mixtures (25 µl) consisted of 1 µl of each set of primers, namely 10 µM 27F/1492R (5'- AGAGTTTGATCCTGGCTCAG-3'/5'- GGYTACCTTGTTAACGACTT -3'), 2.5 µl 10 x Taq Buffer $(NH_4)_2SO_4$, 2.5 µl 2mM dNTPs, 3 µl 25mM $MgCl_2$, 0.5 µl Taq polymerase (1U/µl) and 14.50 µl of ultrapure water. A control reaction mixture containing 1 µl of water instead of *E. coli* was also included in each set of PCR, as a negative control, and a control strain, as a positive

control. All the reagents used were from MBI Fermentas (Vilnius, Lithuania), except for the dNTPs that were from Bioron (Germany).

PCR reactions were performed using a MyCycler Thermal Cycler (Bio-Rad, California, USA). The amplification conditions for total 16S consisted of an initial denaturation at 94 °C for 3 min, followed by 30 amplification cycles consisting of 94 °C for 1 min, 55 °C for 1 min and 72 °C for 2 min, and a final extension step at 72 °C for 10 min. Reaction mixtures that were not immediately used in gel electrophoresis analysis were stored at -20 °C.

After loading the amplification products, 5 µl of each reaction mixture, into a 1.5% (w/v) agarose gel (Lonza, USA), the gels were stained for 10 min with a solution of 0.5 µg mL^{-1} ethidium bromide (Sigma, USA) and visualized under UV light with the imaging Molecular Imager® Gel Doc™ XR+ System (Bio-Rad, USA), just for confirmation of the amplification. PCR products were purified using JETquick PCR Product Purification Spin Kit (GENOMED, Germany) following the manufacturer's instructions. Then, the PCR products were sent to GATC Biotech (Germany) for sequencing (protocol is in appendix B). After the edition of the obtained nucleotide sequences with the help of FinchTV program (Geospiza, USA), the sequences were compared to sequences deposited in the database GenBank using the online Basic Local Alignment Tool BLASTn (NCBI, USA).

5. CRIOPRESERVATION

Pure cultures were stored at -80°C. To this end, each isolate was put to growth in LB broth for 18h and 150 µl of the culture was resuspended in 300 µl of a 45% glycerol solution, prior to immersion in liquid nitrogen. For every isolate was made a replicate.

6. MOLECULAR TYPING AND BOX-PCR

A first approach was done to evaluate which rep-PCR method was more efficient to type this collection of isolates. It was tested the three known sets of primers, namely primers for REP (Repetitive Extragenic Palindromic), ERIC (Enterobacterial Repetitive Intergenic Consensus) and BOX element, for six random and representative isolates (55,58). The protocol used to amplify these repetitive conserved regions was the same as the following procedure (Chapter III; BOX-PCR conditions) with some exceptions for REP and ERIC primers, namely, 1 µl of each primer, namely REP1R/REP2I (5'-IIIICGICGICATCIGGC-3'/5'-NCGICTTATCIGGCCTAC-3') and ERIC1/ERIC2 (5'-AAGTAAGTGACTGGGGTGAGC-3'/5'-ATGTAAGCTCCTGGGGATTCAC-3') was used, and for the REP and ERIC primers the PCR conditions consisted of the same denaturation, amplification and final extension step as the BOX primer, except the annealing temperature, which was 40 °C and 52°C, respectively; better results were obtained with the BOX-PCR method (see in chapter III).

Whole-cell suspensions were prepared from the isolates previously stored at -80°C. For this purpose, 10 µl of each isolate was inoculated in 100 µl of LB broth and grown at 37°C for 8h and 1 µl of each isolate was used as DNA template for the rep-PCR reaction.

The BOX-PCR reaction mixtures (25 µl) consisted of 2 µl of a 10µM BOX A1R primer (5'-CTACGGCAAGGCGACGCTGAC-3'), 6.25 µl NZYTaq 2x Green Master Mix (2.5 mM $MgCl_2$; 200 µM dNTPs; 0.2 U/µl DNA polymerase) (NZYtech, Portugal) and 15.75 µl of ultrapure water. A control reaction mixture containing 1 µl of water instead of *E. coli* was also included in each set of PCR, as a negative control. Two *E. coli* isolates (W33 and W57) were picked and included in every PCR setup as controls to assess the variability in PCR amplification.

PCR reactions were performed using a MyCycler Thermal Cycler (Bio-Rad, California, USA). The amplification conditions consisted of an initial denaturation at 95 °C for 7 min, followed by 30 amplification cycles consisting of 94 °C for 1 min, 53 °C for 1 min and 65 °C for 8 min, and a final extension step at 65 °C for 16 min.

Reaction mixtures that were not immediately used in gel electrophoresis analysis were stored at -20 °C.

7. DNA ELECTROPHORESIS

The amplification products, 5 μl of each reaction mixture, were separated by conventional electrophoresis on a 1.5% (w/v) SeaKem® LE Agarose (Lonza, USA) gel. In all electrophoresis was used a molecular weight marker DNA, GeneRuler™ DNA Ladder Mix (MBI Fermentas, Lithuania). The molecular weight marker was loaded into the two terminal wells and in the middle of the gel as an external reference standard in order to allow the correction of gel irregularities due to electrophoresis process. The gels were run at 80 V, in 1× TAE (5 Prime, Deutschland) for 6 h. At the end of the run the gel was stained for 10 min with a solution of 0.5 μg mL^{-1} ethidium bromide (Sigma, USA) and washed for 1 h at 4 °C. Gel images were captured under UV light with the imaging Molecular Imager® Gel Doc™ XR+ System (Bio-Rad, USA).

8. COMPUTER-ASSISTED BOX-PCR FINGERPRINT ANALYSIS

Gel images were entered into a genomic fingerprint analysis program, GelCompar II® software (Applied Maths, Belgium), in which were normalized and analyzed. The positions of fragments (bands) on each gel were normalized by using the 1-kb ladder from 100 to 10.000 bp as an external reference standard. Normalization with the same set of external standards allowed comparison between independent gels. Fingerprint images were added to a database and compared by performing a statistical analysis.

8.1 STATISTICAL ANALYSIS

Statistical analysis was used to determine the relatedness of DNA fingerprints obtained from molecular typing technique BOX-PCR and to assess the genetic diversity of the *E. coli* library sampled.

8.1.1 DENDROGRAM CONSTRUCTION

Banding patterns were compared using Pearson product-moment correlation coefficient, a densitometric curve-based method that evaluates the intensity as well as the position of the bands to generate pairwise similarity scores that were subsequently used for cluster analysis. Based on individual clusters and the similarity scores between each two positive control strains, Pearson coefficient proved more accuracy for BOX-PCR comparisons than the other methods that account only for band position. For these comparisons, a 1.0% optimization setting was found to give the highest similarity recognition among multiple samples of the control strains.

A dendrogram was generated using the unweighted pair group method using arithmetic mean (UPGMA) constructed by using Pearson's similarity coefficient. This dendrogram included all the isolates of the *E. coli* libraries, in which was applied the settings of the dendrogram construction of the control strains; thus, the similarity value of the cluster that contained all positive control strains served as the similarity cutoff to identify distinctive BOX fingerprints.

8.1.2 SAMPLING SATURATION ANALYSIS

In order to determine if sampling was either or not near saturation, a rarefaction curve was constructed using EcoSim 1.0 (63) software, considering BOX-PCR fingerprints with a cutoff of >85% similar as single strain types, or clonal lines. The EcoSim software provides a computer-sampling algorithm of rarefaction, in which a specified number of individuals are randomly drawn from a community sample. The process is repeated many times to generate a mean and a variance of species diversity (64).

8.1.3 DIVERSITY INDICES

Diversity indices were calculated on the basis of BOX-PCR patterns using the Shannon diversity index and equitability.

Shannon's diversity index was calculated using EcoSim 1.0 (63) software according to:

$$\mathbf{H} = \sum_{i=1}^{s} - (P_i * \ln P_i)$$

Where:

H = Shannon diversity index

P_i = fraction of the entire population made up of species i

S = numbers of species encountered

$\sum$ = sum from species 1 to species S.

On the other hand, equitability index was calculated from H according to equation:

$\mathbf{J} = H/\ln S$

Where:

J = equitability index

S = number of strain types.

III. RESULTS AND DISCUSSION

1. MICROBIOLOGICAL QUALITY OF WATER

Biological parameters were measured in order to assess the microbiological quality of the water of the Berlenga beach. The figure III.1 represents the data obtained for *E. coli* counts within the water collected from the beach.

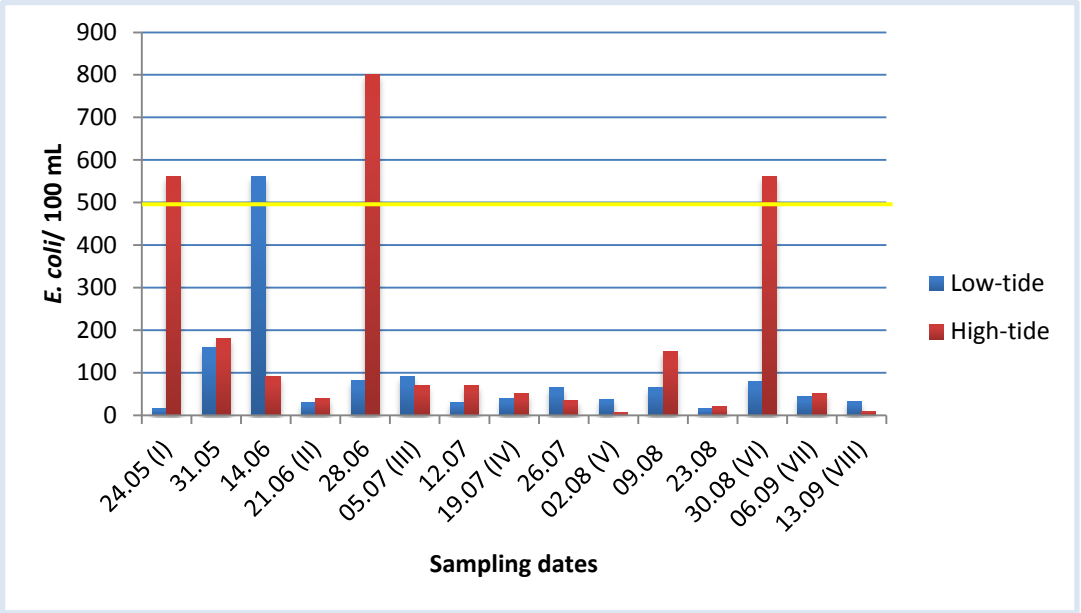


Figure III.1. *E. coli* counting per 100 mL of water collected at low-tide and high-tide moments. Between parentheses (in x axis) are the corresponding main campaigns of water, effluent and feces sampling. The yellow line corresponds to the maximum recommendable/allowable *E. coli* counts value legally permitted (Decree-Law 135/2009).

The analysis of figure III.1 shows that the *E. coli* counts exceeds the limit of the permitted by the Portuguese legislation in some occasions, i.e., values of the counts are outweighed of the MR (maximum recommendable) and MA (maximum allowable) value for coastal waters (52). There are three moments which overcome 500 *E. coli per* 100 mL of water sampled at high-tide, and one at low-tide. These moments occurred in the beginning of the summer season and after in August, when the human

affluence to the island may have been lower given the weather conditions in that period. In fact, the better results for the water quality seem to be when the human affluence to the island is higher. Thus, the microbiological quality of the beach water may be related with the presence of the seagulls, which stay less in the beach when humans are present, and if so this may indicate that the contamination peaks can be directly related to the presence of the seagulls on the beach.

2. *ESCHERICHIA COLI* LIBRARY

The *E. coli* strains were isolated from three distinct sources from the Berlenga Island: beach water, gull feces and human wastewater. After a total of eight campaigns of sampling through summer season, 939 isolates of *E. coli* were obtained. These 939 isolates consisted of 342 isolates from beach water, 427 isolates from gull feces and 170 isolates from effluent presumptively from human origin (table III.1).

Table III.1. *E. coli* isolates sampled and used in the study.

Date	Sampling campaign	Nº. isolates/source			Nº. isolates/sampling campaign
		Water (W)	Feces (F)	Effluent (E)	
23 - 27/May	I	50	49	NP	99
06 - 10/June	NP				
20 - 24/June	II	50	50	0*	100
04 - 08/July	III	50	50	9*	109
18 - 22/July	IV	46	50	50	146
01 - 05/August	V	21	83	22	126
15 - 19/August	NP				
29/August - 02/September	VI	54	63	51	168
05 - 09/September	VII	49	52	20	121
12 - 16/September	VIII	22	30	18	70
Total		**342**	**427**	**170**	**939**

NP, Not performed due to technical and/or meteorological issues.

***, Results due to technical adjustments.**

There were numerous characteristics that were crucial in the choice of *E. coli* as the indicator organism for this study. Namely and mainly, being a ubiquitous intestinal bacterial flora of warm-blooded animals and, once it has been largely used as an indicator of fecal pollution in aquatic environments in most of the MST approaches studies, was a factor taken into account for the selection of the indicator organism once is our main aim in this study to discriminate between human and animal fecal contamination (43,65). In addition to that, this microorganism has other good characteristics that makes it a good candidate as a general indicator of fecal pollution, such as being easily detectable, normally not pathogenic to humans, and is present at concentrations much higher than the pathogens it predicts (34). Thus, the detection of *E. coli* primarily provides evidence of fecal pollution and secondarily reflects the possible presence of bacterial, viral and parasitic enteric pathogens (10,14,18).

In contrast, despite there are other microbial indicators, such as microorganisms belonging to *Enterococcus* genus for example, that also have been used successfully as indicators of fecal pollution and are especially reliable as indicators of health risk in marine environments and recreational waters (34). However, it is known that environmental reservoirs of enterococci exist and that regrowth of these organisms may be possible once they are introduced into the environment (34). Furthermore, human pathogens are not always accompanied by enterococci and vice versa (66).

Diversified phenotypic and genotypic methods have been used in MST studies employing *E. coli* as an indicator, including rep-PCR (33,43). Most of them are library dependent as the present study, requiring a host origin database. There are several studies that used *E. coli* libraries, concluding it is a good candidate for MST approach. Stoeckel and his coworkers (2004) used *E. coli* from eight libraries of host known-source isolates against an unknown-source library to evaluate reproducibility, accuracy, and robustness of seven phenotypic and genotypic MST methods. They concluded that not all the methods were efficient to all the measures they were testing, but in general, PFGE and rep-PCR protocols achieved better results (33).

The genetic heterogeneity and temporal and spatial variability of *E. coli* populations have also been assessed in some previous studies (34,67). Beyond this, in

another study was highlighted the importance of local and temporally specific libraries, even in very small study areas, once they concluded that the ability to match environmental isolates to a host origin database may depend on a significant number of environmental and host origin isolates that ideally are not geographically separated (68). The present study was thought taking into account these questions in the moment of planning the sampling strategy and as it can be seen, the approach was efficient. It was possible to sample the environmental and host sources in a reasonable number, being the larger one the library of isolates collected from the environmental source as it should be; although we had some obstacle in collecting the effluent sample in the beginning due to technical issues or due to weather conditions that not even allowed the boat to leave of the quay (figure III.2).

2.1 INFLUENCE OF LIBRARY SIZE USING *ESCHERICHIA COLI*

In comparison with other studies, the size of the *E. coli* isolates library obtained seems to be reasonable. Several studies have analyzed larger collections of *E. coli* isolates because they encompass many different hosts and environmental samples; however within each type of source the quantity of isolates are very similar as in the present study, or have greater discrepancy between the source groups (57,67,69). However, in overall, in most of the studies, the number of sampling isolates for each type of sample is generally lower and more homogeneous among the source groups analyzed, resulting in smaller libraries, even though few do not have the number of isolates balanced along the different source groups (41,56,60,70–73).

Although there is still controversy concerning which may be the ideal size of the library, it has been suggested that a library size of 20.000 to 40.000 isolates may be the necessary number to capture all the *E. coli* diversity present in the environment, while others defend that the database may require a few hundred isolates per source to the point of representativeness be reached (67,73,74); yet this is not clarified as being an crucial factor to this type of studies, once the studies usually achieve their objectives, apart from the library size. Johnson *et al.* (2004) obtained a considered good average of rate of correct classification to their group sources (82.2%) for the 2,466 rep-PCR DNA fingerprints analyzed; however they

concluded that the increase of the size of the known-source library did not necessarily lead to an increase in the ability to correctly assign strains to the correct source group, once they compared this results with a smaller library obtained from one of their previous studies, and in fact the average rate of correct classification of the larger library was lower in 4.2% (75).

3. *ESCHERICHIA COLI* 16S RRNA SEQUENCING

The edition of the sequences originated by the 16S rRNA gene sequencing enabled the comparison of these sequences with sequences deposited in the GenBank database (NCBI, USA). After the alignment of the sequences with those of the database it was possible to conclude that the isolates that were being tested belong to *E. coli* species, being the homology with the online database sequences of more than 98%.

4. MOLECULAR TYPING

E. coli isolates from the two potential sources of fecal pollution, gulls and sewage, were characterized using BOX-PCR DNA fingerprinting technique and compared with the *E. coli* isolates derived from beach water samples. The analysis of the DNA fingerprints was done in order to assess their genetic variability and to discriminate them according to their source.

The three primers specific to a repetitive sequence highly conserved within the bacteria genome were initially tested with a set of random isolates picked from the three sample types. The figure III.3 is an example of some of the preliminary PCR tests done to decide which rep-PCR method was more suitable for the isolates of this study fingerprinting analysis.

The rep-PCR DNA fingerprinting technique is relatively quick, easy, highly discriminatory and inexpensive to perform and confers the possibility of high-throughput applications, making it an ideal method for MST studies. Thus, provides a

powerful and convenient tool to analyze bacteria diversity (73). As with many of the other PCR typing methods, the results of rep-PCR can be returned in a relatively short amount of time and require a minimum amount of DNA for typing. Based on the method used and the number of repetitive sequences present in the strain, these methods can be highly discriminatory (59).

Rep-PCR targeting the BOX A1R elements of *E. coli* has been evaluated by a number of scientists to distinguish bacterial strains (70,75).

In the present study the fingerprints obtained showed that more complex fingerprint patterns were obtained when amplified with BOX primer. Moreover, some of the isolates that were successfully amplified when the BOX primer was used did not produce reliable fingerprints with neither the REP nor ERIC primers. Consequently, only BOX-derived DNA fingerprints were used in the remainder of the study. Similarly, previous studies concluded that the discriminatory efficacy of BOX-PCR was superior to REP-PCR, in a comparison of the ability of these two methods to discriminate 154 *E. coli* isolates of seven source groups (human, duck, geese, chicken, pig, sheep and cow) (75). Carson *et al.* (71) also reported that rep-PCR DNA fingerprinting done using BOX A1R primers produced a 96.6% average rate of correct classification for human and nonhuman *E. coli* isolates. Still, Ma *et al.* (57) performed rep-PCR with REP, BOX A1R, and $(GTG)_5$ primers to differentiate fecal *E. Coli* isolates from human and nonhuman sources and concluded that the discriminatory efficacy of BOX-PCR was superior to both REP- and $(GTG)_5$-PCR, which allowed the sort of the *E. coli* isolates into the correct source groups (human, cattle, sheep, duck, goose, chicken, and swine).

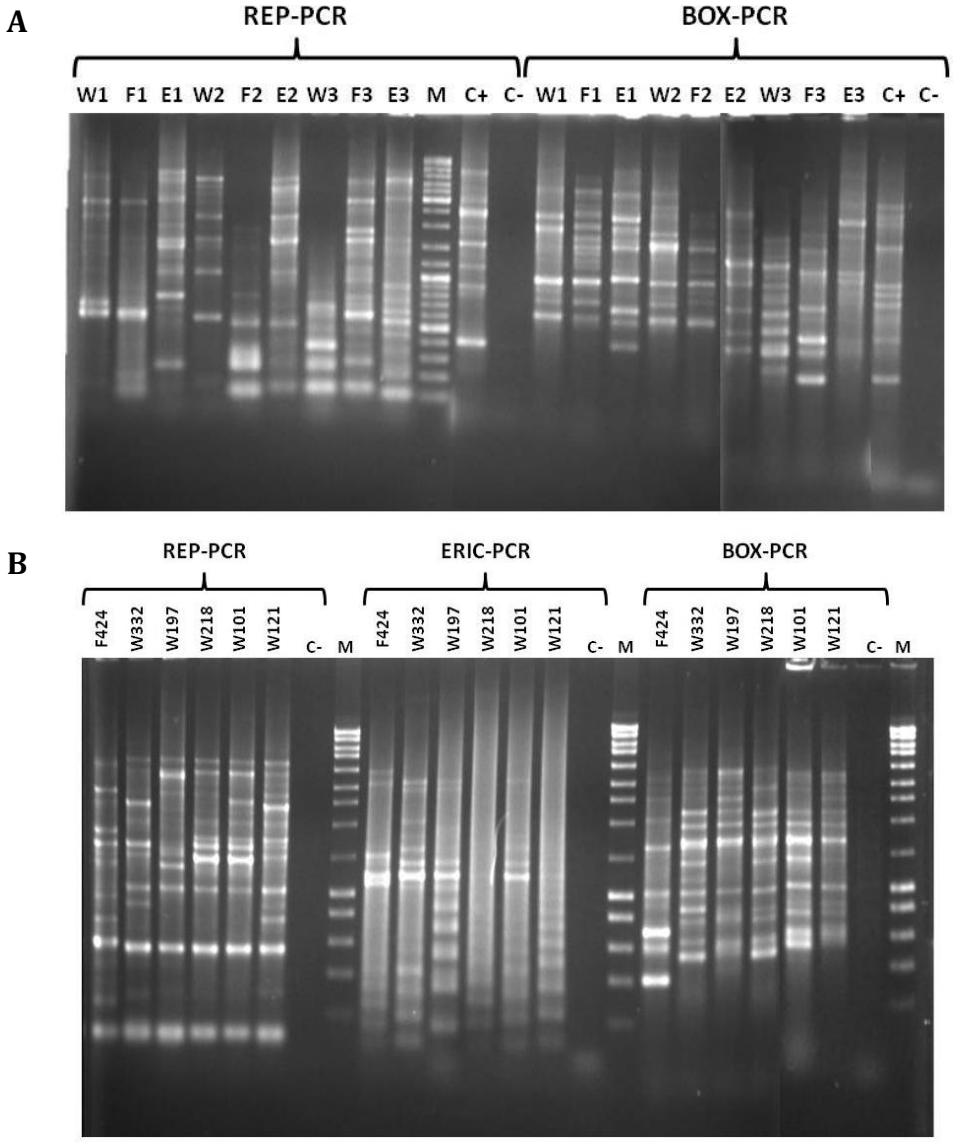


Figure III.3. DNA fingerprints obtained with rep-PCR method for a set of random isolates picked from the collection of *E. coli* isolates, where W is water isolates, F, feces isolates, and E, effluent derived isolates. (A) – BOX and REP-PCR for three different isolates picked randomly of each source; (B) – BOX, REP and ERIC-PCR for six random isolates. The C- corresponds to the negative control of the PCR reactions and the M corresponds to the molecular weight ladder (GeneRuler™ DNA Ladder Mix, MBI Fermentas, Lithuania).

On the other hand, McLellan *et al.* (76) reported a 79.3% average rate of correct classification for *E. coli* isolates analyzed using REP primers. The same study reported that REP-PCR and ERIC-PCR produced comparable, although not identical, results in overall dendrogram groupings similarity indices for 101 selected fecal *E.*

coli isolates from four host groups (human, dog, gull and cattle) (76). In addition, Leung *et al.* (2004) also documented that ERIC-PCR was not an effective tool in distinguishing *E. coli* between animal and human sources (77).

Other approaches may be possible to be implemented in the present study, but once currently none of them stands out from the others as being superior, rep-PCR typing, particularly BOX-PCR seems to be the most framed to this study and moreover, this methodology is already well implemented and optimized in our laboratory, since it has already been performed in previous studies (78–81). Additionally, all of the known MST approaches have disadvantages. There are numerous studies that have subjected comparisons between two or more different MST methodologies. For instance, Price and his coworkers (2007) concluded that PFGE performs better than source classification using ARA (82). Additionally, developing DNA PFGE data for scat sample isolates and water sample analysis can be significantly more expensive (three times more) and time consuming than developing ARA and in addition to the cost differential, the PFGE methodology requires about 10 times the amount of time to implement than ARA (82). Similarly, with other studies Parveen and his coworkers found that PFGE profiles analysis of *E. coli* isolates do not stood out in differentiating between human and nonhuman isolates, since according to them this method detects small differences on a sequence that may not be related with a specific bacterial characteristic, such as host source (43); the main differences are investigator dependent as the results highlighted the need to modify and optimize analytical and statistical methods in order to minimize sources of error (83).

Moreover, a stepwise combination of rapid screening methods, and detailed source tracking techniques may produce higher discrimination between closely related strains, although, this resolution may not be necessary since adequate discrimination can be achieved with a single-primer approach, as it can be seen in this study; and in the future then form the basis for future management of environmental water quality including improved microbial risk assessment.

Determining which method or combination of methods to use for any given situation will depend on a number of factors including: specific question to be answered, detail required to answer the question (i.e. broad scale results –

human/non-human versus detailed results – human, livestock species, wildlife species), availability of resources (cost of analysis varies depending on technique used, and size of the water-body), time constraints, and ability to access a lab or facilities with expertise to analyze the samples.

4.1 BOX-PCR

A total of 939 isolates were used as templates for PCR performed with the BOX A1R primer (58). Complex fingerprint patterns were obtained for approximately 97% of the isolates studied. The strains that did not produce fingerprints were excluded from the analysis. As a result, the analyses were performed using 926 isolates, from which 169 strains were from human effluent, 423 strains from gull fecal samples and 334 strains from the water.

The amplicons sizes of the PCR products ranged from 300 to 8.000 bp. Individual lanes generally contained from 20 to 25 PCR product bands, although for some *E. coli* isolates this number of bands was superior (figure III.4) (all images of the gels can be found in appendix C).

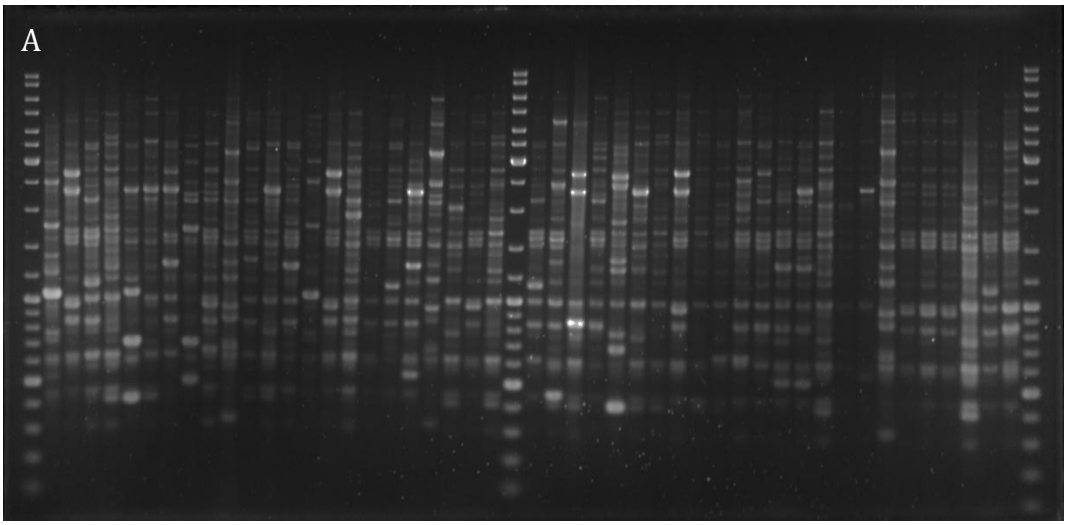


Figure III.4. Examples of BOX-PCR DNA fingerprint patterns of *E. coli* strains obtained from water (A), feces (B) and effluent (E) (continued on next page). The lanes of the ends and center contained an external standard, a 1-kb molecular weight ladder (GeneRuler™ DNA Ladder Mix, MBI Fermentas, Lithuania).

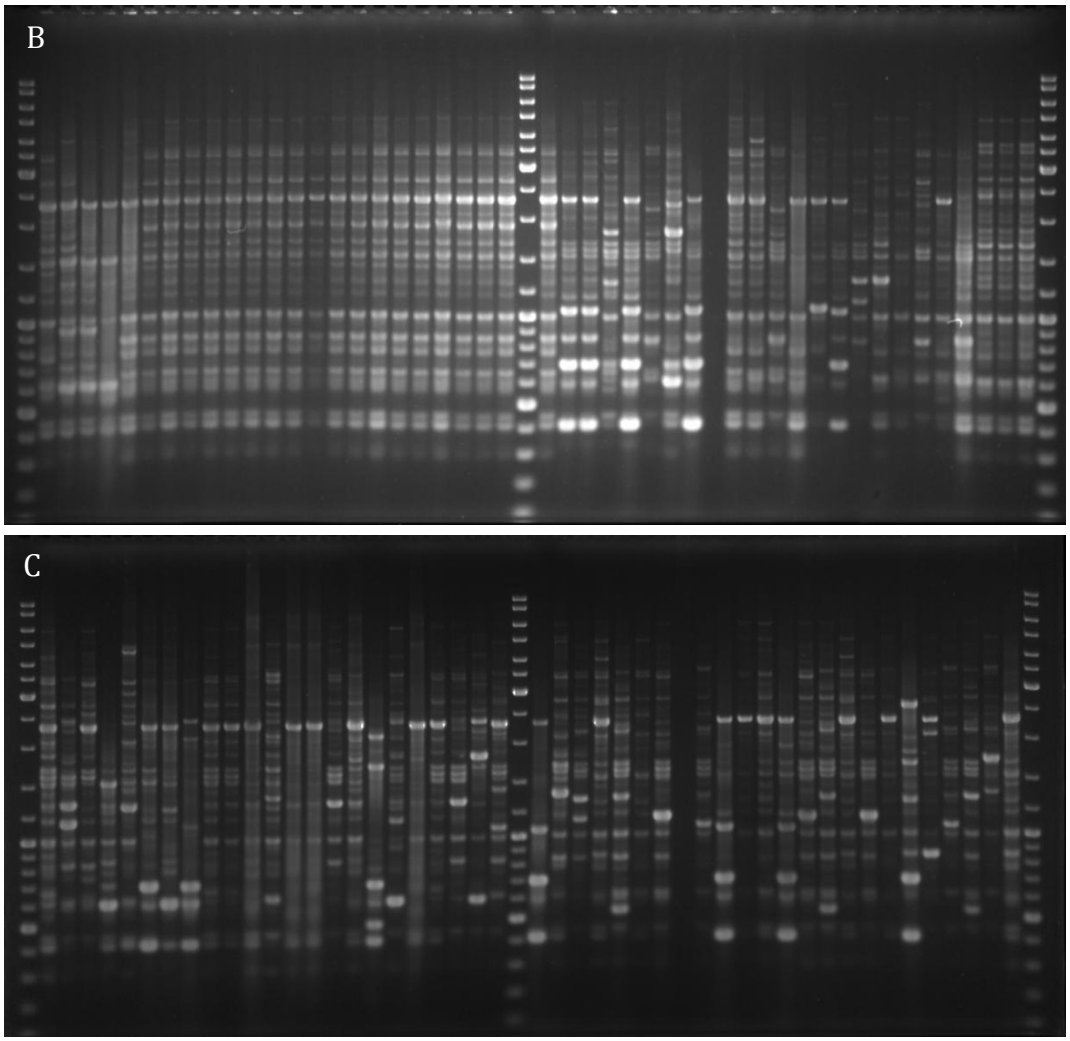


Figure III.4. (Cont.) Examples of BOX-PCR DNA fingerprint patterns of *E. coli* strains obtained from water (A), feces (B) and effluent (E). The lanes of the ends and center contained an external standard, a 1-kb molecular weight ladder (GeneRuler™ DNA Ladder Mix, MBI Fermentas, Lithuania).

4.2 REPEATABILITY OF BOX-PCR METHOD

The reproducibility of each BOX-PCR fingerprinting method was examined using the fingerprints of two *E. coli* strains used as reference controls (W33 and W57), which were included in all the PCR assays and gel runs experiments in order to assess the bias of PCR and gel-to-gel variation. DNA fingerprint patterns assembled from all individual PCRs, each of which run on a separate agarose gel, when analyzed with the Pearson's correlation coefficient had an average similarity for most of the replicas above 85%, for both *E. coli* reference controls (figure III.5).

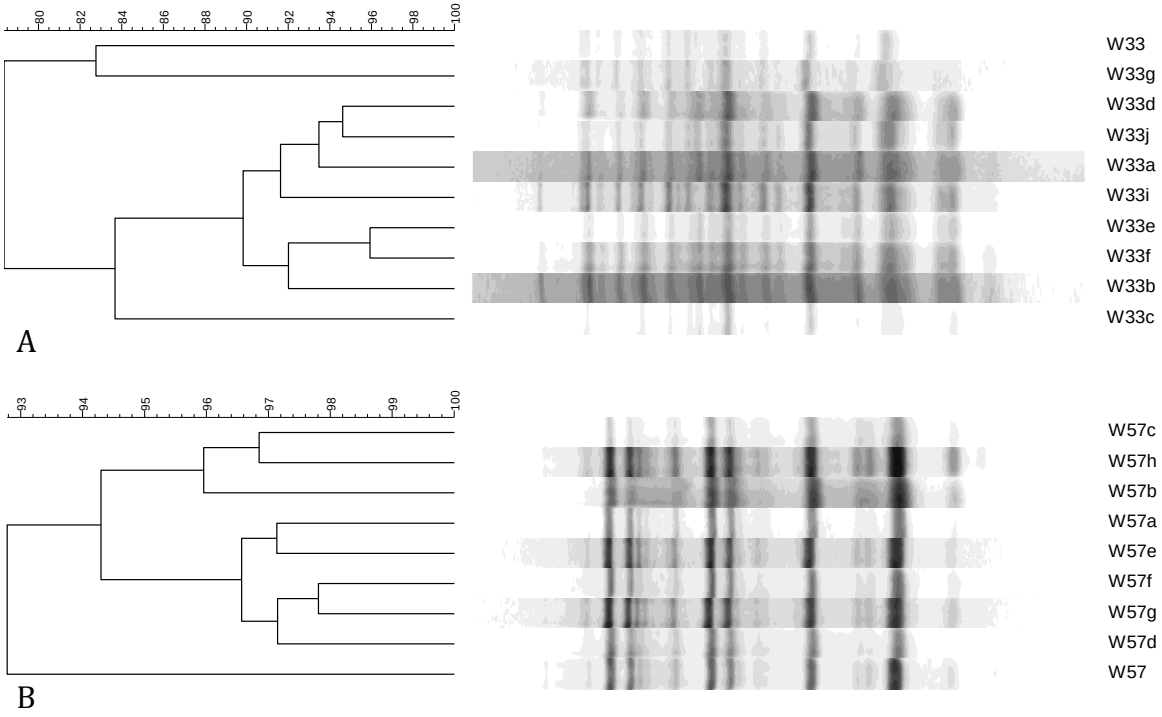


Figure III.5. Comparison of DNA fingerprint patterns of the two reference *E. coli* strains generated by BOX-PCR. (A) Fingerprints were generated using *E. coli* W33 (water) isolate as DNA template. (B) Fingerprints were generated using *E. coli* W57 (water) isolate as DNA template.

This indicates that, for this type of data, the Pearson's product-moment correlation coefficient with UPGMA method of tree building was superior to Jaccard's coefficient for higher rates of similarity within the control isolates replicas (data not shown). This is similar to results reported by Häne *et al.* (84), who demonstrated that for complex DNA fingerprints, such as those produced in this study, a curve-based method such as Pearson's product-moment correlation coefficient was consistently more reliable to identify similar or identical DNA fingerprints in comparison to band matching formulas, such as Jaccard coefficient. Similarly, Louws *et al.* (85) reported that curve-based statistical methods worked best for analysis of complex banding profiles generated by rep-PCR, since comparison of curve data is less dependent on DNA concentration in loaded samples and is relatively insensitive to background differences in gels; and also, Albert *et al.* (86) performed a statistical evaluation of rep-PCR DNA fingerprint data and reported that Pearson's product-moment coefficient had ability to correctly classify fingerprints of 584 *E. coli* isolates.

5. DENDROGRAM ANALYSIS

To determine the relatedness of strains a dendrogram based on BOX-derived fingerprint data was constructed by using Pearson's correlation coefficient and the UPGMA method. The analysis of the composite dendrogram obtained including all host and environmental strains (n = 926) did not reveal distinct grouping of strains according to host source, but rather into multiple closely related subclusters apparently host-specific, which is consistent with the findings of other studies (global dendrogram is presented in appendix D).

Hagedorn and coworkers were able to classify fecal streptococci isolates into host and environmental groups (humans, dairy cattle, beef cattle, chickens, deer, and waterfowl) using antibiotic resistance patterns; however, the initial protocol of five antibiotics did not provide satisfactory separation of isolates from known sources by which some overlap occurred between the human and nonhuman clusters (74).

Another study, of Dombek *et al.* (2000) showed that although the dendrogram analysis may have been useful for separating isolates into human and nonhuman source groups, the isolates were clearly closely related, resulting in some major clusters of human isolates mixed with some isolates waterfowl-derived (75).

Therefore, in some instances, it may be sufficient to identify unknown environmental *E. coli* isolates to the level of larger groupings, rather than to the level of strain types.

5.1 SIMILARITY CUTOFF

In order to simplify the clusters analysis a similarity score value of 85% was used as a cutoff for designating strain types; this value was based upon comparison of patterns generated by repeated analysis of the reference strain W57 (n = 9), where the similarity value of the cluster that contained all this *E. coli* strains served as a similarity cutoff of 90% to identify distinctive identical patterns. This strain was used as a control in all PCR runs and loaded on every gel, independently. For this reason, wild-type strains with similarity scores above 85% were considered, with a high degree of certainty, the same strain type. Clusters with a similarity coefficient lower than the cutoff value were considered distinctive BOX fingerprints. This strategy resulted in a high diversity in fingerprints, as the 926 isolates yielded 314 distinctive strain types.

5.2 WATER ISOLATES AFFILIATION WITH HOST-SOURCES

In the light of the data obtained from the cluster analysis, major divisions in the dendrogram between 50 and 55% of similarity were analyzed (table III.2). Of the 10 main groups formed, 5 of them revealed a prominent presence of gull isolates against effluent isolates; corresponding to 88.3, 81.4, 100, 66.4 and 78.6% of fecal gull isolates of the total number of isolates in the clusters (cluster 1, 2, 3, 6 and 9, respectively); one of which is solely grouped with water and feces isolates (cluster 3). In contrast, there was only a group that showed equality between the presence of fecal gull and effluent isolates (cluster 10), and another one that only grouped two

effluent isolates without any water affiliated isolates (cluster 8). In addition, there were four groups that had somewhat half feces and effluent isolates (clusters 4, 5, 7 and 10), leading to what was considered as an indeterminate group for the presumptive source. Therefore, of the 302 water isolates, a total of 230 affiliated with feces clusters, corresponding to more than 75% of the water isolates.

Table III.2. Percentage of water isolates affiliated with feces and effluent, in which only clusters composed mainly of a source were considered for affiliation (> 65%). The number of clades were based on the isolates groups originated by cutting between 50-55% of similarity on the complete dendrogram (*n* = 926).

No. of clades	**Feces**		**Effluent**		**Predominant source**	**Water**	
	No. of isolates	Percentage of isolates	No. of isolates	Percentage of isolates		No. of isolates *per* clade	Percentage of affiliation with source
1	68	88.3	9	11.7	Feces	62	**76.2 (Feces)**
2	79	81.4	18	18.6	Feces	79	
3	15	100	0		Feces	1	
4	30	57.7	22	42.3	Not determined	40	**0 (Effluent)**
5	27	57.4	20	42.6	Not determined	25	
6	97	66.4	49	33.6	Feces	61	
7	13	56.5	10	43.5	Not determined	4	**23.8 (Unknown)**
8	0		2	100	Effluent	0	
9	66	78.6	18	21.4	Feces	27	
10	8	50	8	50	Not determined	3	

This lower percentage of effluent isolates affiliation with the environmental isolates may, in part, be due to the smaller number of fingerprints analyzed for this category, once the library sizes of host origin isolates are limited (normally consisting of 35 to about 500 isolates) making broader comparisons to larger populations of *E. Coli* in the environment difficult (41,43,67,71,72,75,76).

In order to counterpoise the number of isolates of the two potential fecal sources and to eliminate the possible bias associated with the use of libraries with different sizes, the isolates fecal-originated from campaigns I, II, III and V were

removed, maintaining though all the water isolates. Therefore, a dendrogram was constructed with similar conditions of the previous one but using 196 feces profiles, 138 effluent profiles and 333 water profiles. Results are shown in table III.3.

Cluster analysis of this group of isolates above the 50% of similarity resulted in a total of 20 divisions of the dendrogram. Hereupon, between 50 and 55% of similarity, only 7 of them had significance to the analysis once these were the ones that shown water affiliation of the isolates of known-sources with water. Four clades grouped mainly feces isolates, corresponding to 66.1, 79.2 and 100% (two of them) (clades number 3, 4, 5 and 7, respectively). And the remaining three clustered as much feces as effluent, such as clusters 1, 2 and 6, corresponding to the indeterminate group for the presumptive source. This corresponded to water affiliation to seagull feces clusters of 61.7%, and the rest to the not determined source grouping.

Table III.3. Percentage of water isolates affiliated with number of feces and effluent isolates balanced, (*n* = 196 and *n* = 138, respectively). Only clusters composed mainly of a source were considered for affiliation (> 65%). The number of clades were based on the isolates groups originated by cutting between 50-55% of similarity on the complete dendrogram (*n* = 667).

No. of clades	**Feces**		**Effluent**		**Predominant source**	**Water**	
	No. of isolates	Percentage of isolates	No. of isolates	Percentage of isolates		No. of isolates *per* clade	Percentage of affiliation with source
1	30	46.9	34	53.1	Not determined	55	**61.7 (Feces)**
2	15	45.5	18	54.5	Not determined	39	
3	39	66.1	20	33.9	Feces	88	
4	19	79.2	5	20.8	Feces	44	**0 (Effluent)**
5	2	100	0		Feces	12	
6	12	57.1	9	42.9	Not determined	4	**38.3 (Unknown)**
7	21	100	0		Feces	14	

This reduction of the isolates derived from the two host types in order to balance their number in the library seems to continue to have gull feces isolates more prevalent over the isolates effluent-derived, as well as within the total dendrogram major groups, as shown in table III.2.

5.3 FECES AND EFFLUENT DENDROGRAM ANALYSIS

In order to verify the results above, and evaluate if the differentiation between the isolates solely host-derived was distinctive, a minor dendrogram of the BOX fingerprints was constructed using the same conditions used previously (UPGMA method and Pearson coefficient with 1% of optimization) (dendrogram presented in appendix D).

Cluster analysis of the BOX fingerprints derived from the gull and human sources (n = 334) did not produced significant clusters of each host-specific sources, i.e. the overall arrangement of the sub-clusters across the dendrogram was not by host group but intermixed. Despite this, the analysis of the dendrogram taking into account the cutoff by the 85% of similarity demonstrated sub-clusters of closely related strains with high rates of similarity (between 90 and 98% of similarity values); and additionally, an evaluation of the clusters formed with a 50 to 55% similarity range, demonstrated major clusters produced always dominated by isolates derived from the gull feces or from the effluent. This corresponded to 11 major clades dominated by feces isolates and in turn, 8 major clades dominated by effluent isolates (data not shown).

6. DIVERSITY AND RELATIVE ABUNDANCE OF *E. COLI* STRAINS

The diversity among strains isolated from the feces host source were bigger than from the strains isolated from the water environmental samples and effluent samples (table III.4). As shown in the following table, 62.7% of the isolates obtained from gull feces, 58.4% of the isolates from water and 56.5% of the effluent isolates correspond to unique strains in the data set.

In overall, this corresponds to an average of 59.2% of unique BOX fingerprints obtained from the known and unknown-sources sampling.

Table III.4. Diversity and relative abundance of *E. coli* strains of the collection.

Strain type/ clade size (no. of isolates)	**Water isolates (*n* = 334)**		**Feces isolates (*n* = 423)**		**Effluent isolates (*n* = 169)**	
	No. of clades	Percentage of isolates in category	No. of clades	Percentage of isolates in category	No. of clades	Percentage of isolates in category
1[a]	97	58.4	113	62.7	39	56.5
2	36	21.7	17	9.4	12	17.4
3	13	7.8	23	12.8	3	4.3
4	3	1.8	8	4.4	2	2.9
5	7	4.2	3	1.7	5	6.2
6	4	2.4	5	2.8	3	4.3
7	1	0.6	2	1.1	1	1.5
8	1	0.6	4	2.2	1	1.5
9	2	1.2	1	0.6	2	2.9
10	1	0.6	1	0.6		
>10	1 (12)	0.6	1 (14)	0.6	1 (13)	1.5
			1 (19)	0.6		
			1 (32)	0.6		
Total no. of strain types	166		180			69

[a] Only one isolate found with a given rep-PCR fingerprint pattern

[b] A strain type was defined as a set of isolates with more than 85% similarity based on comparison of BOX-PCR fingerprint patterns using the Pearson coefficient.

6.1 SAMPLING SATURATION ASSESSMENT

The sampling saturation was assessed in order to evaluate whether *E. coli* obtained were sufficient to capture the genetic diversity present within the *E. coli* populations sampled, i.e. the representativeness of the *E. coli* collections (figure III.7). To the rarefaction curve construction BOX-PCR fingerprints with 85% or greater similarities (based on Pearson coefficient, UPGMA and 1% optimization) were considered the same strain type, corresponding in general, to either a difference of one band between strains or a difference in the intensity of the amplified bands. The rarefaction curve was constructed by summing the number of genotypes (species richness) that accumulated with the successive addition of isolates (abundance of isolates).

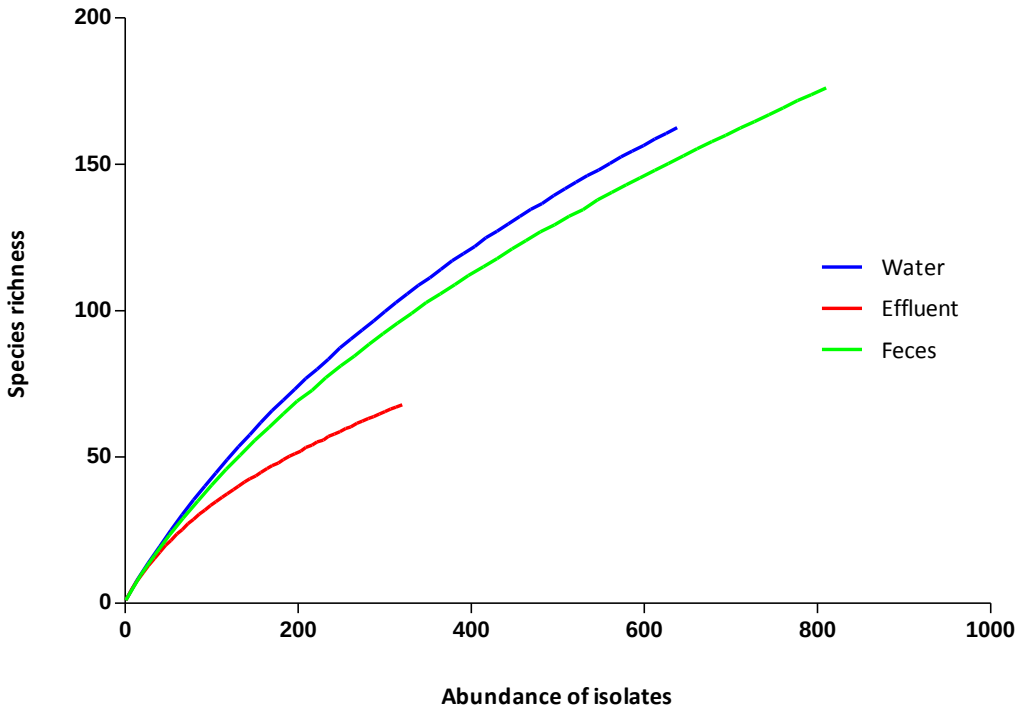


Figure III.7. Rarefaction curve generated for assessment of sampling saturation of *E. coli* strains collection by determination of the number of strain types (species richness) found in each group (water, gulls and effluent) for the number of isolates (abundance of isolates) sampled.

Despite the library size of 926 DNA fingerprints, genetic diversity has not been saturated. This is evidenced by the apparent shape of the rarefaction curve that did not become asymptotic to neither one of the isolates groups, between isolates numbers (sampling effort) and accumulation of new strain types. Moreover, and in accordance to this, within the host-source groups, more than half of the genotypes occurred only once in the database (59.6%), and a limited number occurred multiple times (table III.2). This may indicate that sampling efforts were not enough, and there is a lot more of *E. coli* diversity in the populations sampled.

Neither one of the isolates groups, effluent, feces and water appeared to be near sampling saturation for possible strains. The average slope of the line was the highest for the water isolates with an average of 0.24 (e.g. 24 unique strains per 100 sampled); though the gull isolates had a line slope very close, corresponding to an average slope of 0.20. These last, despite having a higher number of isolates sampled appear to be far from sampling saturation. In contrast, the effluent seems to be the one with a lower slope average (0.19), being the only group that appears to have a good representation of possible strains.

Despite these average slope values seem to be scarce to obtain a good library representativeness, some studies indicate that the lower value obtained for their study was for beach water with 0.26 of average slope of the line (which corresponds to the higher value of the present study) where the authors consider a reasonable representation of the collection (69).

Since our rarefaction curve did not become asymptotic, our data cannot be used to predict the ultimate size that our fingerprint library needs to be.

Taking into account the library size that has been suggested to capture all the genetic diversity present in *E. coli* populations, the data show that the use of relatively small libraries, that may not take into account the tremendous genetic diversity present in *E. coli* (76), will make broader comparisons to larger populations of these organisms in the environment difficult.

Among the strategies suggested to avoid this underrepresentation problem in the literature, the present study made the possible efforts to accomplish them. Namely, developing moderate-sized libraries for a confined geographical area, in order to obtain isolates only from the animals present in the study area, only

sampling from the animals pertinent to the study site, and those likely to have an impact on the targeted water-body needing to be examined (73). However, even if a careful planning was made prior to the scientific project (SEAGULL) beginning, in this case, the animals sources were unpredictable by the way that could vary over time, depending either directly on the weather conditions, in the case of human-source (touristic activities), and indirectly, in the case of the seagulls present in the island (human presence and search for food at the sea or on the shore).

6.2 INDICES

Diversity indices for each source group were calculated based on the corresponding number of strain types and are shown in table III.5. In the case of water isolates library, even though the number of strain types is relatively lower than for feces, the obtained data for Shannon diversity index shows a greater value, as well as for the equitability. As for the isolates collected from the effluent, the data set show a lower value for the diversity index, despite the greater equitability value.

Table III.5. Diversity of *E. coli* based on the DNA fingerprinting patterns of strains recovered from feces, effluent and water samples.

Strains sources	Parameters		
	No. of strain types	Shannon diversity index (*H*)	Equitability (*J*)
Water	166	3.10	0.606
Feces	180	3.05	0.587
Effluent	69	2.63	0.621

Thus, the results of the equitability index reflect complex populations of *E. coli* but that are dominated by a small number of strain types.

IV. CONCLUSIONS

Fecal pollution is considered as a worldwide concern for public health. This type of contamination can be a problem in coastal waters environments which are associated with wildlife and human populations, as the area of the present study, the Berlenga Island. Such entities can introduce fecal pollution that not only degrades water quality, but also restricts its use for harvesting seafood and recreational activities.

An MST approach was tested using isolates of *E. coli*, that is a common environmental bacterium used widely as a specific indicator of fecal pollution in water environments. The rep-PCR methodology used, namely BOX-PCR fingerprinting method, was chosen in order to evaluate the effectiveness in determining the source of the *E. coli* isolates sampled. Our results suggest that the molecular typing of the isolates collected using BOX A1R primer was useful to differentiate between different *E. coli* strains of human and animal origin. Therefore, the major aim of this study was achieved, given the analysis carried out during the study; our results suggest that gull feces may be the dominant source of the water contamination detected in the beach water. Despite this, we cannot exclude the contribution of other sources of pollution, such as human.

Thus, BOX-PCR methodology for MST can be recommended as a powerful approach to be adopted in future similar studies, since it allowed the differentiation of the origin of fecal pollution, between the two hosts studied, seagulls and humans.

With this study was also possible to construct a library of isolates derived from the three types of samples (water, effluent and feces), which is an important condition for further studies. However, although in future works it may not be as important as for this study, the results suggest the samples did not reach the saturation point at the sampling moment.

Moreover, the biological parameters tested in the water samples every week demonstrated a higher probability of water contamination in moments when humans are less present in the island. This may mean that the summer season is not the moment most representative of this marine seabird population problem, but may be the rest of the year.

Although extensive field testing is required to determine the efficacy of these assays and much larger referencing databases must be accumulated before these methods could be used for routine natural environmental monitoring, these assays appear to provide promising diagnostic tools for tracking non-point sources of fecal pollution.

V. REFERENCES

1. Amado A, Gafeira C, Teixeira A, Preto A. Plano de ordenamento da reserva natural das berlengas - Relatório para discussão pública. 2007.

2. Queiroga H, Leão F, Coutinho M. Candidatura das Berlengas a Reserva da Biosfera da UNESCO. 2008.

3. The International Coordinating Council. 18 new Biosphere Reserves added to UNESCO's Man and the Biosphere (MAB) Programme. 2011.

4. Resolução do Conselho de Ministros n.º 180/2008. Diário da República; 1.ª série - N.º 228; 2008.

5. Cúrdia J, Carvalho S, Ravara A, Gage J, Rodrigues A, Quintino V. Deep macrobenthic communities from Nazaré Submarine Canyon (NW Portugal). Scientia Marina. 2004;68:171–80.

6. Pardal M. Zooplankton biomass, abundance and diversity in a shelf area of Portugal (the Berlenga Marine Natural Reserve). Life and Marine Sciences. 2001;18A:25–33.

7. Vidal E, Medail F, Tatoni T. Is the yellow-legged gull a superabundantbird species in the Mediterranean? Impacton fauna and flora, conservation measuresand research priorities. Biodiversity and Conservation. 1998;7:1013–26.

8. Maranhão P, Rodrigues NV, Oliveira P, Alberto J. Guia de espécies submarinas, Portugal - Berlengas. 2008.

9. World Health Organization. Water, sanitation, and hygiene links to health: Facts and figures. 2004;

10. Touron A, Berthe T, Gargala G, Fournier M, Ratajczak M, Servais P, et al. Assessment of faecal contamination and the relationship between pathogens and faecal bacterial indicators in an estuarine environment (Seine, France). Marine pollution bulletin. 2007 Sep;54(9):1441–50.

11. Pond K. Water recreation and disease: plausibility of associated infections: acute effects, sequelae, and mortality. IWA Publishing; 2005.

12. World Health Organization Sustainable Development and Healthy Environments. Bathing Water Quality and Human Health: Fecal Pollution. 2001;

13. World Health Organization. Guidelines for Safe Recreational Water Environments: Coastal and Fresh Waters. 2003;1.

14. World Health Organization Europe. Outbreaks of waterborne diseases. 2009.

15. Dufour A, Snozzi M, Koster W, Bartram J, Ronchi E, Fewtrell L. Assessing microbial safety of drinking water: Improving approaches and methods. IWA Publishing; 2003.

16. Soller J a, Schoen ME, Bartrand T, Ravenscroft JE, Ashbolt NJ. Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination. Water research. 2010 Sep;44(16):4674–91.

17. Santo Domingo JW, Edge TA. Identification of primary sources of faecal pollution. In: Safe Management of Shellfish and Harvest Waters. 2010.

18. Field KG, Samadpour M. Fecal source tracking, the indicator paradigm, and managing water quality. Water research. 2007 Aug;41(16):3517–38.

19. Tyagi V, Chopra A, Kazmi A, Kumar A. Alternative microbial indicators of faecal pollution: Current perspective. Iranian journal of environmental health science and engineering. 2006;3(3):205–16.

20. Kay D, Bartram J, Prüss A, Ashbolt N, Wyer MD, Fleisher JM, et al. Derivation of numerical values for the World Health Organization guidelines for recreational waters. Water research. 2004 Mar;38(5):1296–304.

21. Mansilha CR, Coelho CA, Heitor AM, Amado J, Martins JP, Gameiro P. Bathing waters: new directive, new standards, new quality approach. Marine pollution bulletin. 2009 Oct;58(10):1562–5.

22. Figueras MJ, Polo F, Inza I, Guarro J. Past, Present and Future Perspectives of the EU Bathing Water Directive. Marine pollution bulletin. 1997 Mar;34(3):148–56.

23. Rees G, Pond K, Johal K, Pedley S, Rickards A. Microbiological analysis of selected coastal bathing waters in the UK, Greece, Italy and Spain. Water research. 1998;32(8):2335–40.

24. European Comission. Council Directive 91/271/EEC of 21 May 1991 concerning urban waste water treatment. Official Journal of the European Communities. 1991;

25. Roccaro P, Mancini G. Water intended for human consumption - Part I : Compliance with European water quality standards. Desalination. 2005;176:1–11.

26. European Commission. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. Official Journal of the European Communities. 1998;

27. Council of the European Union, European Parliament. Directive 2006/7/EC of the European Parliament and of the council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. 2006;

28. European Parliament, Council of the European Union. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. 2000;

29. United States Environmental Protection Agency. Microbial Source Tracking Guide Document. 2005;

30. Meays CL, Broersma K, Nordin R, Mazumder A. Source tracking fecal bacteria in water: a critical review of current methods. Journal of environmental management. 2004 Oct;73(1):71–9.

31. Griffith JF, Weisberg SB, McGee CD. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. Journal of water and health. 2003 Dec;1(4):141–51.

32. Stewart JR, Ellender RD, Gooch J a, Jiang S, Myoda SP, Weisberg SB. Recommendations for microbial source tracking: lessons from a methods comparison study. Journal of water and health. 2003 Dec;1(4):225–31.

33. Stoeckel DM, Mathes MV, Hyer KE, Hagedorn C, Kator H, Lukasik J, et al. Comparison of seven protocols to identify fecal contamination sources using Escherichia coli. Environmental science & technology. 2004 Nov 15;38(22):6109–17.

34. Scott T, Rose J, Jenkins T. Microbial Source Tracking: current methodology and future directions. Applied and environmental microbiology. 2002;68(12):5796–803.

35. Santo Domingo JW, Bambic DG, Edge T a, Wuertz S. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. Water research. 2007 Aug;41(16):3539–52.

36. Barrell R, Hunter P, Nichols G. Microbiological standards for water and their relationship to health risk. Communicable disease and public health. 2000 Mar;3(1):8–13.

37. Bitton G. Microbial indicators of fecal contamination: application to microbial source tracking. 2005.

38. Nnane DE, Ebdon JE, Taylor HD. Integrated analysis of water quality parameters for cost-effective faecal pollution management in river catchments. Water research. 2011 Mar;45(6):2235–46.

39. Roslev P, Bukh AS. State of the art molecular markers for fecal pollution source tracking in water. Applied microbiology and biotechnology. 2011 Mar;89(5):1341–55.

40. Sinton L, Finlay R, Hannah D. New Zealand Journal of Distinguishing human from animal faecal contamination in water: a review. New Zealand journal of marine and freshwater research. 1998;32(2):323–48.

41. Guan S, Xu R, Chen S, Odumeru J, Gyles C. Development of a procedure for discriminating among Escherichia coli isolates from animal and human sources. Applied and environmental microbiology. 2002;68(6):2690–8.

42. Yang Z, Wu X, Li T, Li M, Zhong Y, Liu Y, et al. Epidemiological survey and analysis on an outbreak of gastroenteritis due to water contamination. Biomedical and environmental sciences. 2011;24(3):275–83.

43. Parveen S, Hodge NC, Stall RE, Farrah SR, Tamplin ML. Phenotypic and genotypic characterization of human and nonhuman Escherichia coli. Water research. 2001 Feb;35(2):379–86.

44. Bosch A. Human enteric viruses in the water environment: a minireview. International microbiology: the official journal of the Spanish Society for Microbiology. 1998 Sep;1(3):191–6.

45. Simpson JM, Santo Domingo JW, Reasoner DJ. Microbial Source Tracking: state of the science. Environmental science & technology. 2002;36(24):5279–88.

46. Santo Domingo JW, Ashbolt NJ. Fecal pollution of water. Encyclopedia of earth. 2010;

47. Ashbolt NJ. Microbial contamination of drinking water and disease outcomes in developing regions. Toxicology. 2004 May 20;1(3):229–38.

48. Figueras MJ, Borrego JJ. New perspectives in monitoring drinking water microbial quality. International journal of environmental research and public health. 2010 Dec;7(12):4179–202.

49. Council of the European Communities. Council Directive of 8 December 1975 concerning the quality of bathing water (76/160/EC). Official Journal of the European Communities.

50. European Parliament and the council of the European Union. Directive 2006/7/EC of the European Parliament and of the council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC. 2006.

51. Holland PG. Water is life: the water framework directive. 2002;

52. Abelho M. Manual de monitorização microbiológica ambiental Manual de monitorização. 2010;

53. Cabral JPS. Water microbiology. Bacterial pathogens and water. International journal of environmental research and public health. 2010 Oct;7(10):3657–703.

54. Xie M, Yin H-qun, Liu Y, Liu J, Liu X-duan. Repetitive sequence based polymerase chain reaction to differentiate close bacteria strains in acidic sites. Transactions of nonferrous metals society of China. 2008 Dec;18(6):1392–7.

55. Versalovic J, Koeuth T, Lupski JR. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic acids research. 1991 Dec 25;19(24):6823–31.

56. Carlos C, Alexandrino F, Stoppe NC, Sato MIZ, Ottoboni LMM. Use of Escherichia coli BOX-PCR fingerprints to identify sources of fecal contamination of water bodies in the State of São Paulo, Brazil. Journal of environmental management. 2012 Jan;93(1):38–43.

57. Ma H-J, Fu L-L, Li J-R. Differentiation of fecal Escherichia coli from human, livestock, and poultry sources by rep-PCR DNA fingerprinting on the shellfish culture area of East China Sea. Current microbiology. 2011 May;62(5):1423–30.

58. Versalovic J, Schneider M, De Bruijn FJ, Lupski JR. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. Methods in molecular and cellular biology. 1994;5(1):25–40.

59. Foley SL, Lynne AM, Nayak R. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2009 Jul;9(4):430–40.

60. Hassan WM, Wang SY, Ellender RD. Methods to increase fidelity of repetitive extragenic palindromic PCR fingerprint-based bacterial source tracking efforts. Applied and environmental microbiology. 2005;71(1):512–8.

61. Ashbolt NJ, Grabow WOK, Snozzi M. Indicators of microbial water quality. In: Water Quality: Guidelines, Standards and Health. 2001.

62. American Public Health Association, American Water Works Association, Water Environment Federation. Standard methods for the examination of water and wastewater. In: American Public Health Association, editor. Washington, DC: 1999.

63. Entsminger G. EcoSim Professional: Null modeling software for ecologists [Internet]. 2012;Available from: http://www.garyentsminger.com/ecosim/index.htm

64. Kesey-Bear. EcoSim 5.0 Help System; Species Diversity [Internet]. Acquired Intelligence Inc. 2000 [cited 2012 Jul 3];Available from: http://esapubs.org/archive/ecol/E081/022/EcoSim Help/Species/Species Diversity.htm

65. Wicki M, Karabulut F, Auckenthaler A, Felleisen R, Tanner M, Baumgartner A. Identification of faecal input sites in spring water by selection and genotyping of

multiresistant Escherichia coli. Applied and environmental microbiology. 2011 Sep 30;

66. Santiago-Rodríguez TM, Dávila C, González J, Bonilla N, Marcos P, Urdaneta M, et al. Characterization of Enterococcus faecalis-infecting phages (enterophages) as markers of human fecal pollution in recreational waters. Water research. 2010 Sep;44(16):4716–25.

67. Hartel PG, Summer JD, Hill JL, Collins JV, Entry JA, Segars WI. Geographic variability of Escherichia coli ribotypes from animals in Idaho and Georgia. Journal of environmental quality. 2002;31(4):1273–8.

68. Kelsey RH, Webster LF, Kenny DJ, Stewart JR, Scott GI. Spatial and temporal variability of ribotyping results at a small watershed in South Carolina. Water research. 2008 Apr;42(8-9):2220–8.

69. McLellan S. Genetic diversity of Escherichia coli isolated from urban rivers and beach water. Applied and environmental microbiology. 2004;70(8):4658–65.

70. Mohapatra BR, Broersma K, Mazumder A. Comparison of five rep-PCR genomic fingerprinting methods for differentiation of fecal Escherichia coli from humans, poultry and wild birds. FEMS microbiology letters. 2007 Dec;277(1):98–106.

71. Carson CA, Shear BL, Ellersieck MR, Schnell JD. Comparison of ribotyping and repetitive extragenic palindromic-PCR for identification of fecal Escherichia coli from humans and animals. Applied and environmental microbiology. 2003;69(3):1836–9.

72. Seurinck S, Verstraete W, Siciliano SD. Use of 16S-23S rRNA intergenic spacer region PCR and repetitive extragenic palindromic PCR analyses of Escherichia coli isolates to identify nonpoint fecal sources. Applied and environmental microbiology. 2003;69(8):4942–50.

73. Johnson LK, Brown MB, Carruthers EA, Ferguson JA, Dombek PE, Sadowsky MJ. Sample size, library composition, and genotypic diversity among natural populations of Escherichia coli from different animals influence accuracy of determining sources of fecal pollution. Applied and environmental microbiology. 2004;70(8):4478–85.

74. Hagedorn C, Robinson S, Filtz J. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. Applied and environmental microbiology. 1999;

75. Dombek PE, Johnson LK, Zimmerley ST, Sadowsky MJ. Use of repetitive DNA sequences and the PCR To differentiate Escherichia coli isolates from human and animal sources. Applied and environmental microbiology. 2000 Jun;66(6):2572–7.

76. McLellan S, Daniels A, Salmore A. Genetic characterization of Escherichia coli populations from host sources of fecal pollution by using DNA fingerprinting. Applied and environmental microbiology. 2003;69(5):2587–94.

77. Leung KT, Mackereth R, Tien Y-C, Topp E. A comparison of AFLP and ERIC-PCR analyses for discriminating Escherichia coli from cattle, pig and human sources. FEMS microbiology ecology. 2004 Jan 1;47(1):111–9.

78. Alves A, Henriques I, Fragoeiro S, Santos C, Phillips AJL, Correia A. Applicability of rep-PCR genomic fingerprinting to molecular discrimination of members of the genera Phaeoacremonium and Phaeomoniella. Plant Pathology. 2004 Oct;53(5):629–34.

79. Alves A, Phillips AJL, Henriques I, Correia A. Rapid differentiation of species of Botryosphaeriaceae by PCR fingerprinting. Research in microbiology. 2007 Mar;158(2):112–21.

80. Alves A, Santos O, Henriques I, Correia A. Evaluation of methods for molecular typing and identification of members of the genus Brevibacterium and other related species. FEMS microbiology letters. 2002 Aug 6;213(2):205–11.

81. Tacão M, Alves A, Saavedra MJ, Correia A. BOX-PCR is an adequate tool for typing Aeromonas spp. Antonie van Leeuwenhoek. 2005 Aug;88(2):173–9.

82. Price B, Venso E, Frana M, Greenberg J, Ware A. A comparison of ARA and DNA data for microbial source tracking based on source-classification models developed using classification trees. Water research. 2007 Aug;41(16):3575–84.

83. Myoda SP, Carson CA, Fuhrmann JJ, Hahm B-K, Hartel PG, Yampara-Lquise H, et al. Comparison of genotypic-based microbial source tracking methods requiring a host origin database. Journal of water and health. 2003 Dec;1(4):167–80.

84. Häne B, Jäger K, Drexler H. The Pearson product-moment correlation coefficient is better suited for identification of DNA fingerprint profiles than band matching algorithms. Electrophoresis. 1993;14(10):967–72.

85. Louws F, Rademaker J, de Bruijn F. The three Ds of PCR-based genomic analysis of Phytobacteria: Diversity, Detection, and Disease Diagnosis. Annual review of phytopathology. 1999 Jan 28;37:81–125.

86. Albert JM, Munakata-Marr J, Tenorio L, Siegrist RL. Statistical evaluation of bacterial source tracking data obtained by rep-PCR DNA fingerprinting of Escherichia coli. Environmental science & technology. 2003 Oct;37(20):4554–60.

87. Faculdade de Ciências e Tecnologia - Universidade Nova de Lisboa. Reservas em Portugal [Internet]. [cited 2012 May 21];Available from: http://campus.fct.unl.pt/afr/ipa_9899/grupo0017_natureza/escrita.htm

VI. APPENDICES

APPENDIX A – CULTURE MEDIA AND REAGENTS

6.3 A.1 CULTURE MEDIA

For all culture media the composition is provided for volumes of 1 liter. All of them were acquired from Merck (Darmstadt, Germany).

A.1.1 Chromocult® Coliform Agar medium:

Composition:

3.0 g Peptone

5.0 g Sodium chloride

2.2 g Sodium di-hydrogen phosphate

2.7 g di-Sodium hydrogen phosphate

1.0 g Sodium pyruvate

1.0 g Tryptophan

10.0 g Agar-agar

1.0 g Sorbitol

70.15 g Tergitol

0.4 g Chromogenic mixture (0.2 g 6-chloro-3-indoxyl-beta-D-galactopyranoside; 0.1 g isopropyl-beta-D-thiogalactopyranoside; 0.1 g 5-bromo-4-chloro-3-indoxyl-D-glucuronic acid)

(pH: 6.8)

Preparation:

1) The broth powder in weight of 26.5 g was dissolved in 1 L of distilled water;

2) Heat in the microwave, by stirring regularly until total dissolution;

3) The medium was cooled and poured into petri dishes;

4) After solidification the plates were kept at temperature of 4°C.

A.1.2 MacConkey Agar medium:

Composition:

17.0 g Peptone from gelatin

1.5 g Peptone from casein

1.5 g Peptone from meat

5.0 g Sodium chloride

10.0 g Lactose

1.5 g Bile salt mixture

0.03 g Neutral red

0.001 g Crystal violet

13.5 g Agar-agar

(pH: 6.8)

Preparation:

1) The broth powder in weight of 50 g was dissolved in 1 L of distilled water;

2) Heat in the microwave, by stirring regularly until total dissolution;

3) The medium was cooled and poured into petri dishes;

4) After solidification the plates were kept at temperature of 4°C.

A.1.3 mFC medium:

Composition:

5.0 g Proteose-peptone

10.0 g Tryptose

3.0 g Yeast extract

5.0 g Sodium chloride

1.5 g Bile salt mixture

12.5 g Lactose

0.1 g Methyl blue (formerly aniline blue)

15.0 g Agar-agar

(pH: 7.4)

Addition: 10mL of a 1% solution of rosolic acid in 0.2N NaOH

Preparation:

1) The broth powder in weight of 52 g was dissolved in 1 L of distilled water;
2) Heat in the microwave, by stirring regularly until total dissolution;
3) The medium was cooled and poured into petri dishes;
4) After solidification the plates were kept at temperature of 4°C.

A.1.4 TSB (Tryptic Soy Broth) medium:

Composition:

17.0 g Peptone from casein
3.0 g Peptone from soymeal
5.0 g Sodium Chloride
2.5 g di-Potassium hydrogen phosphate
2.5 g Glucose
(pH: 7.3)

Note: For TSA solid culture medium, formula may be supplemented with 14 g agar-agar.

Preparation:

1) The broth powder in weight of 30 g was dissolved in 1 L of distilled water;
2) 15 g of microbiological agar was added to the solution;
3) The solution was mixed for dissolution and autoclaved in 121°C for 15 min.;
4) The sterilized medium was cooled and poured into petri dishes;
5) After solidification the plates were kept at temperature of 4°C.

A.1.5 LB (Luria-Bertani) medium:

Composition:

10.0 g Peptone from casein
5.0 g Yeast extract
10.0 g Sodium chloride
(pH 7.0)

Preparation:

1) The broth powder in weight of 25 g was dissolved in 1 L of distilled water;
2) The solution was mixed for dissolution and autoclaved in 121°C for 15 min.;
3) The sterilized medium was cooled and poured into petri dishes;
4) After solidification the plates were kept at temperature of 4°C.

6.4 A.2 REAGENTS AND SOLUTIONS

Composition of some reagents or solutions of general use is described below.

A.2.1 50× TAE (Tris-Acetate-EDTA) buffer (5 Prime, Deutschland):

2 M Tris-Acetate

0.05 M EDTA

(pH 8.3)

A.2.2 6× Loading Dye (MBI Fermentas, Lithuania):

10 mM Tris-HCl (pH 7.6)

0.03% bromophenol blue

0.03% xylene cyanol FF

60% glycerol

60 mM EDTA

APPENDIX B – PCR PRODUCT PURIFICATION PROTOCOL

1. Add 400 µl of solution H1 to the PCR volume assay and mix thoroughly.
2. Place a JETQUICK spin column into a 2 ml receiver tube. Load the mixture from step 1 into the prepared spin column.
3. Centrifuge at >12,000 x g for 1 min. Discard the flowthrough.
4. Re-insert the spin column into the empty receiver tube and add 500 µl of reconstituted solution H2.
5. Centrifuge at >12,000 x g for 1 min.
6. Centrifuge again at the maximum velocity for 1 min.
7. Place the JETQUICK spin column into a new 1.5 ml microfuge tube and add 30 µl of sterile water directly (previously heated at 65 °C) onto the center of the silica matrix of the JETQUICK spin column.
8. Leave at room temperature for 1 min.
9. Centrifuge at >12,000 x g for 2 min.
10. Store the microfuge tubes at 20°C, until use.

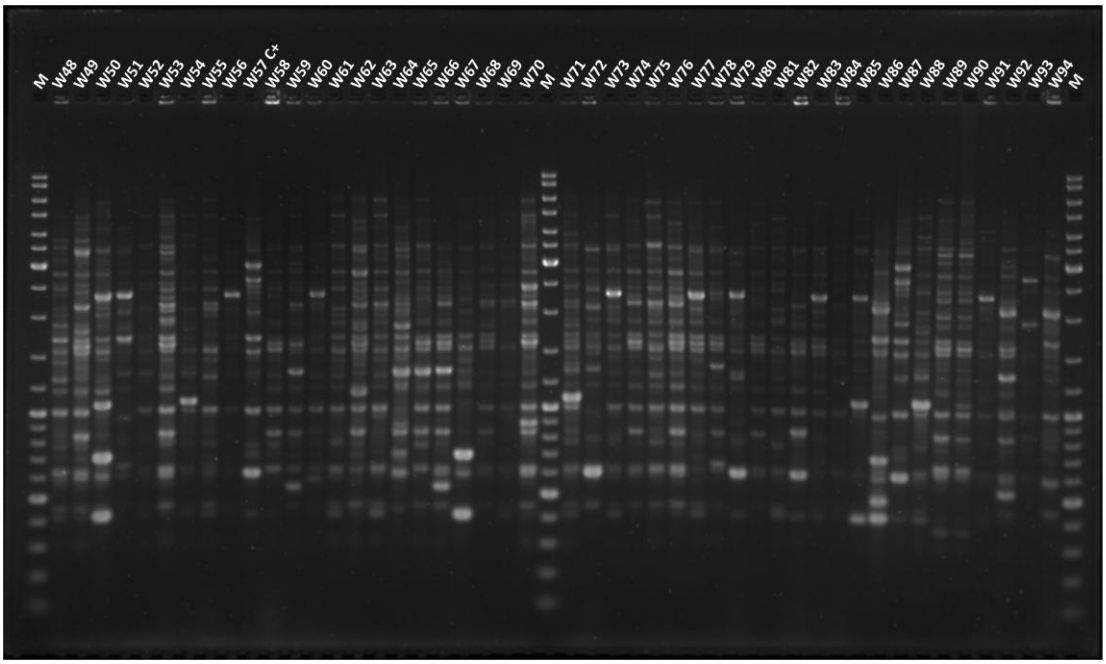
APPENDIX C – BOX-PCR FINGERPRINTS GEL IMAGES

Images of the gels obtained for each BOX-PCR done for all the *E. coli* isolates of the collection. Each PCR was loaded on a 100 wells agarose gel. For this reason, for each PCR loaded into a gel there are two images, one from the top and one from the bottom of the same electrophoresis.

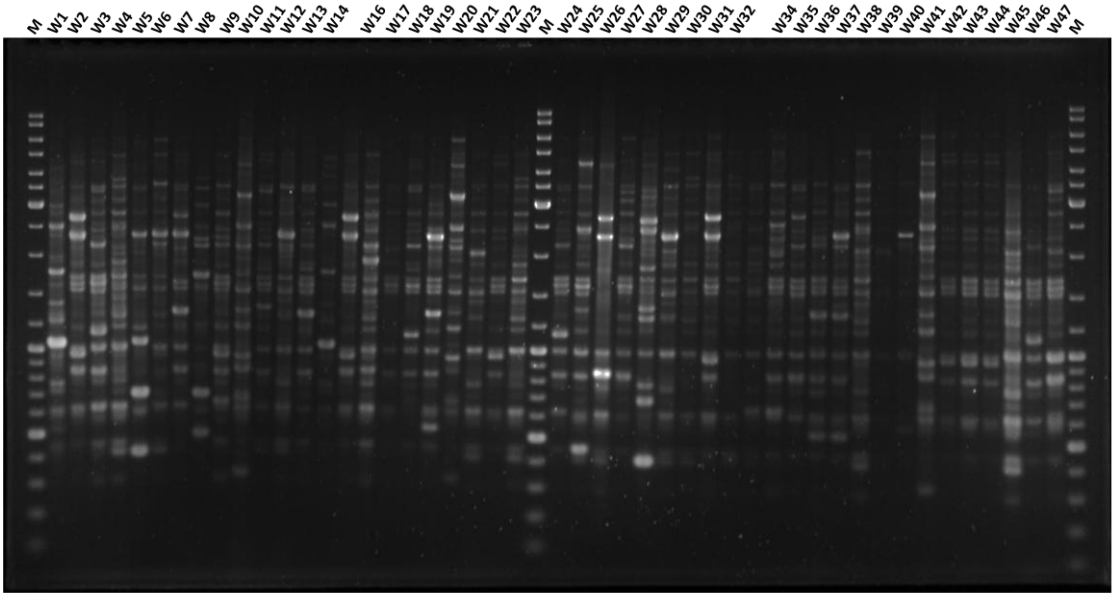
The isolates from water, effluent and gull feces are designated by the W, E and F letters, respectively, followed by the corresponding number, for each type of sample, in the collection; The M corresponds to the molecular weight ladder (GeneRuler™ DNA Ladder Mix, MBI Fermentas, Lithuania).

BOX-PCR 1:

- Top

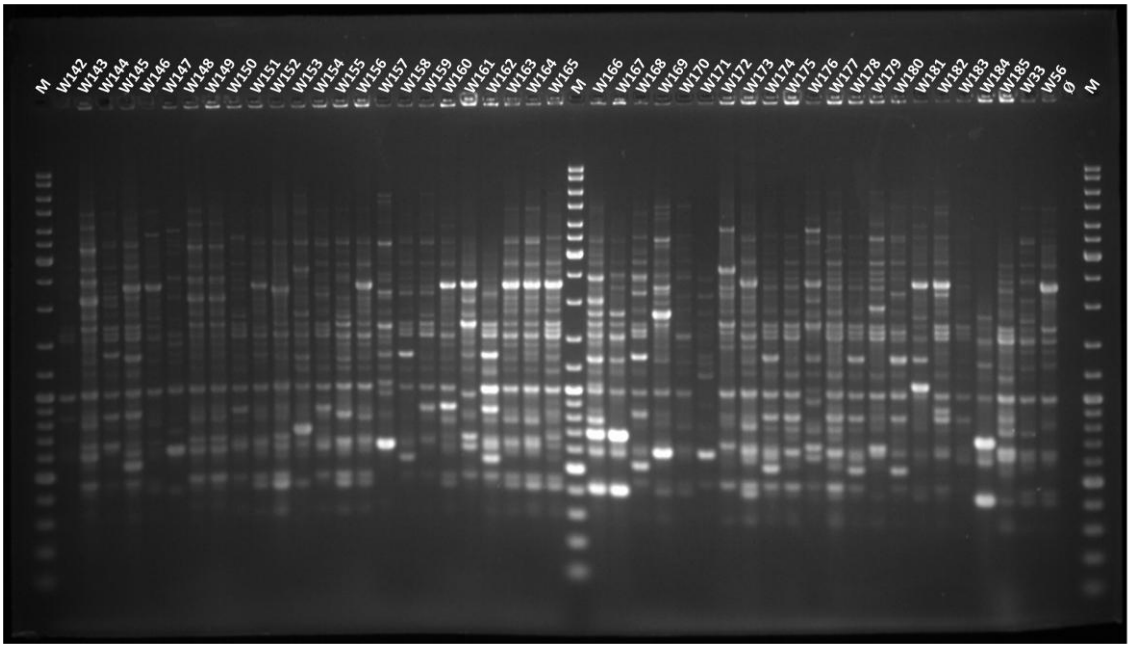


- Bottom

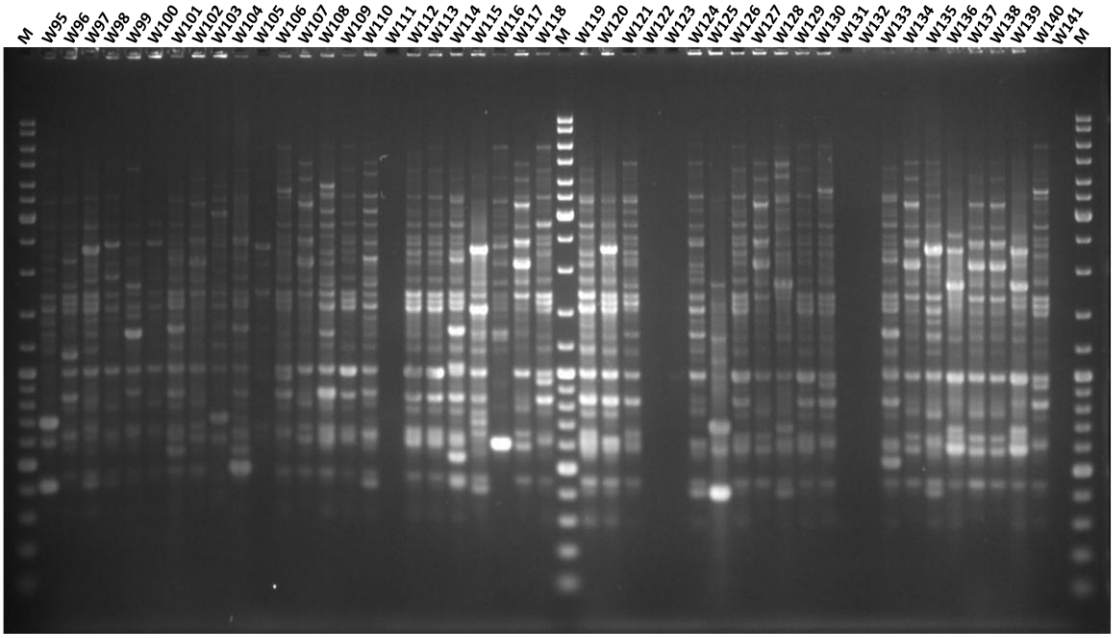


BOX-PCR 2:

- Top

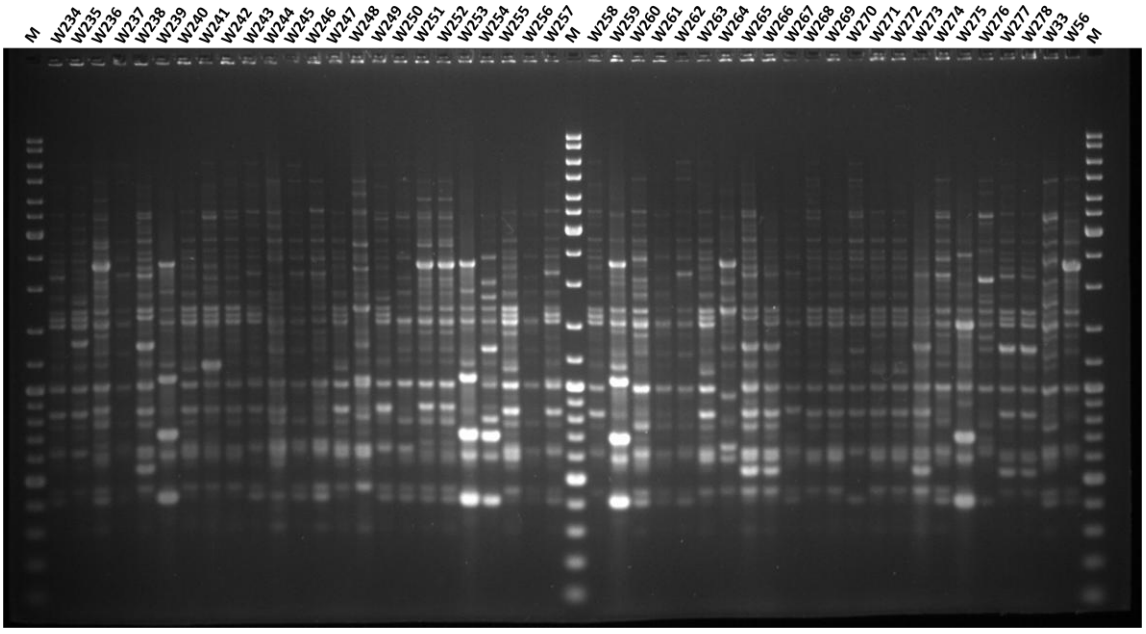


- Bottom

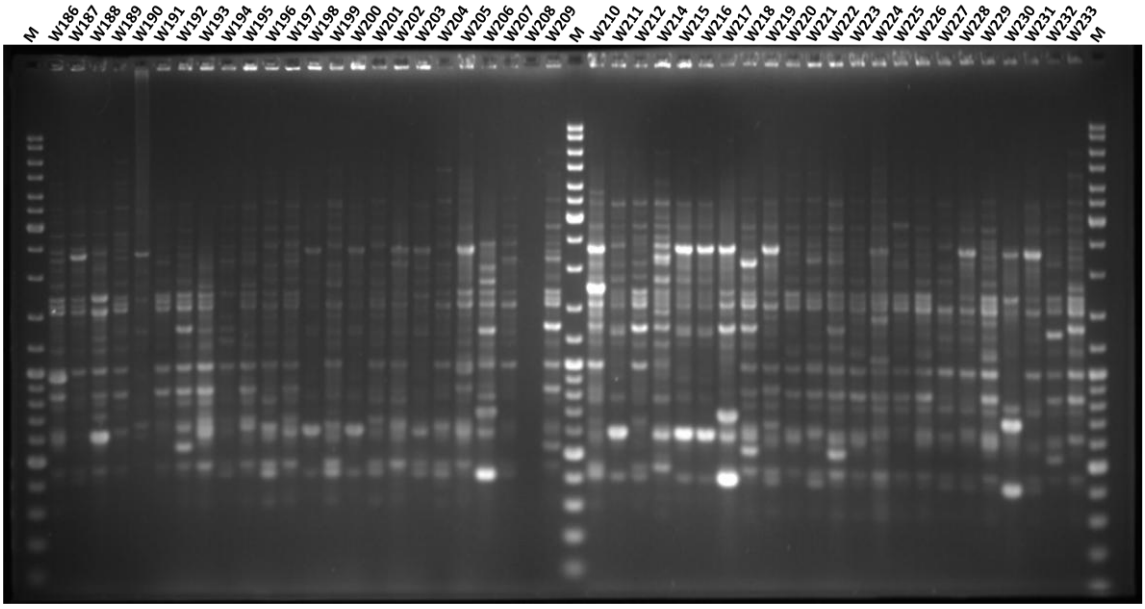


BOX-PCR 3:

- Top

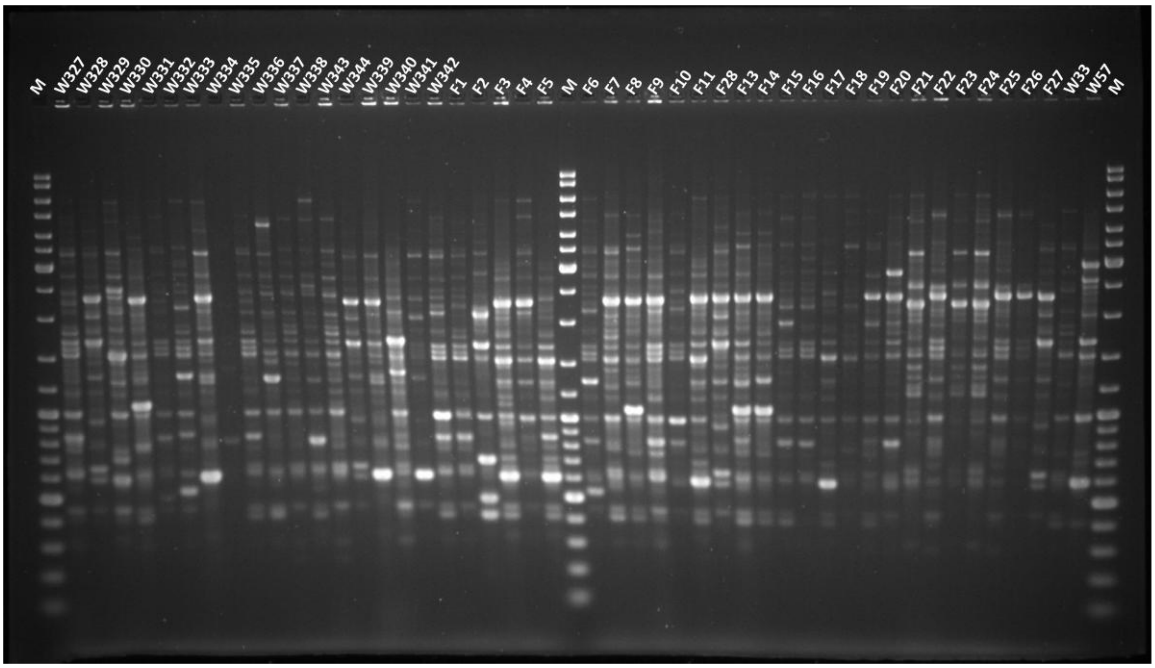


- Bottom

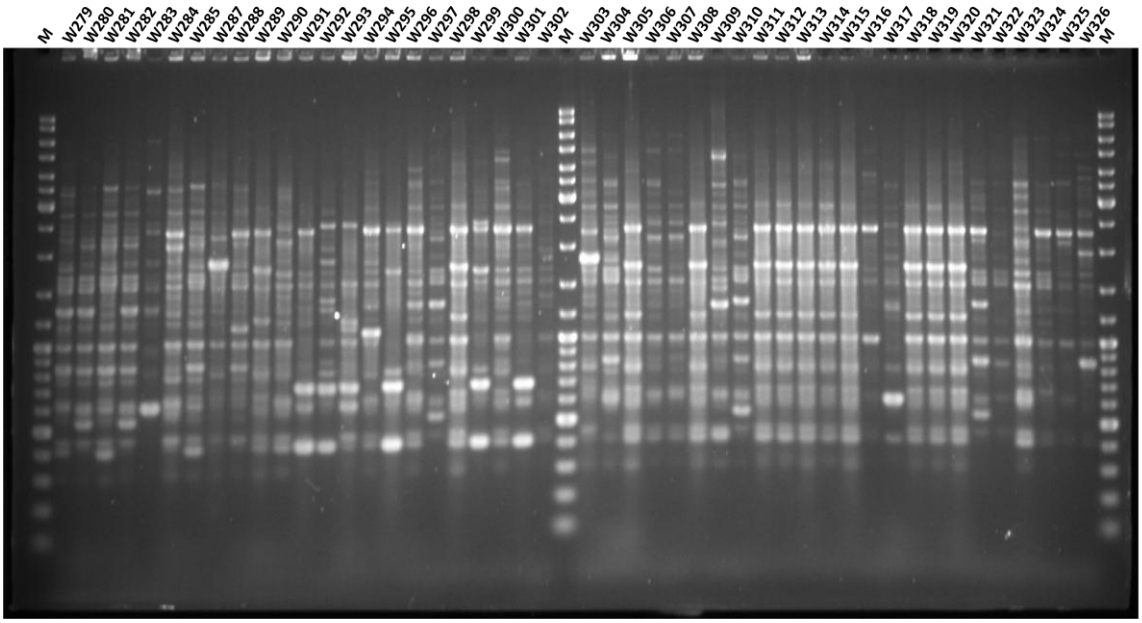


BOX-PCR 4:

- Top

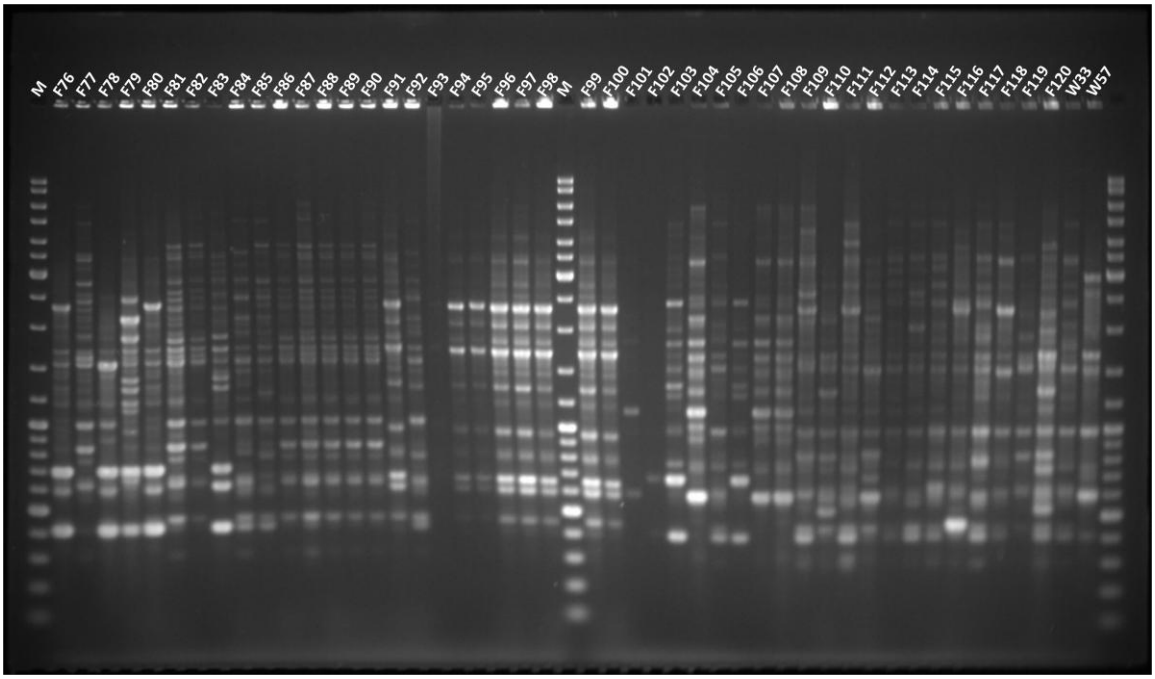


- Bottom

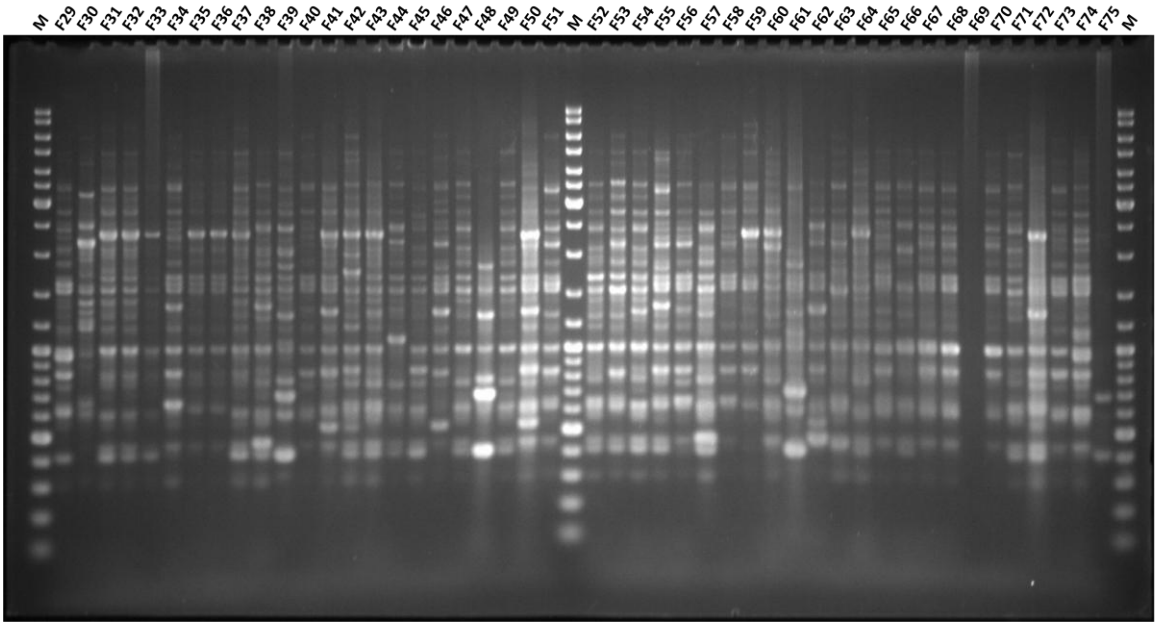


BOX-PCR 5:

- Top

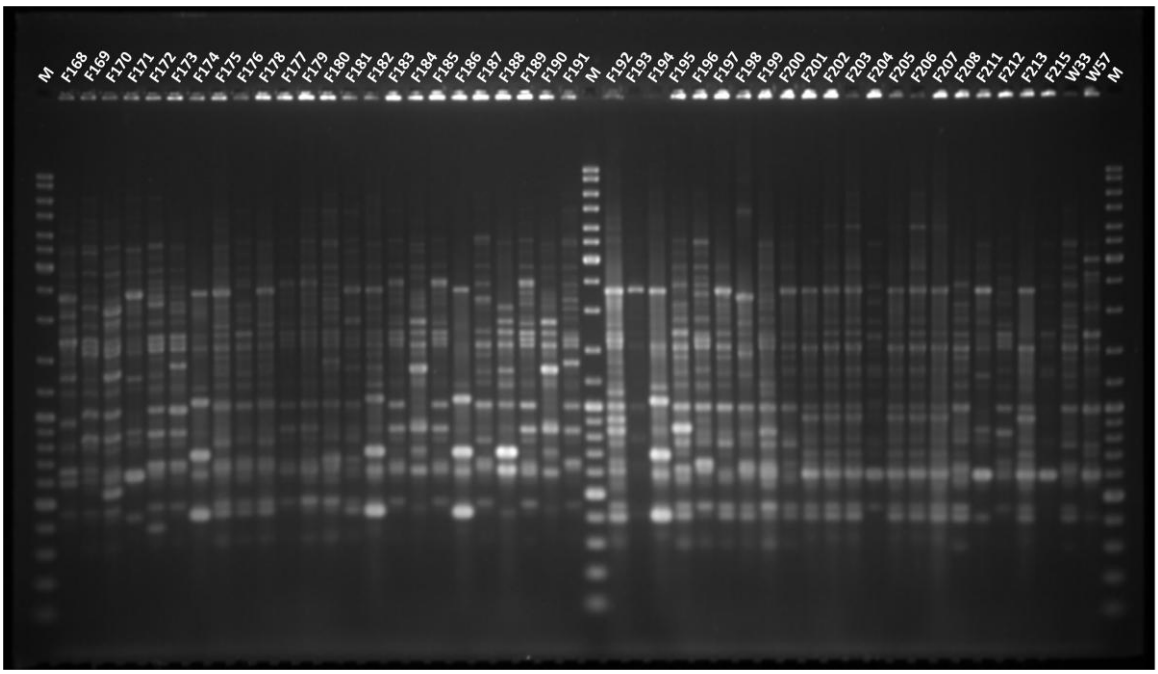


- Bottom

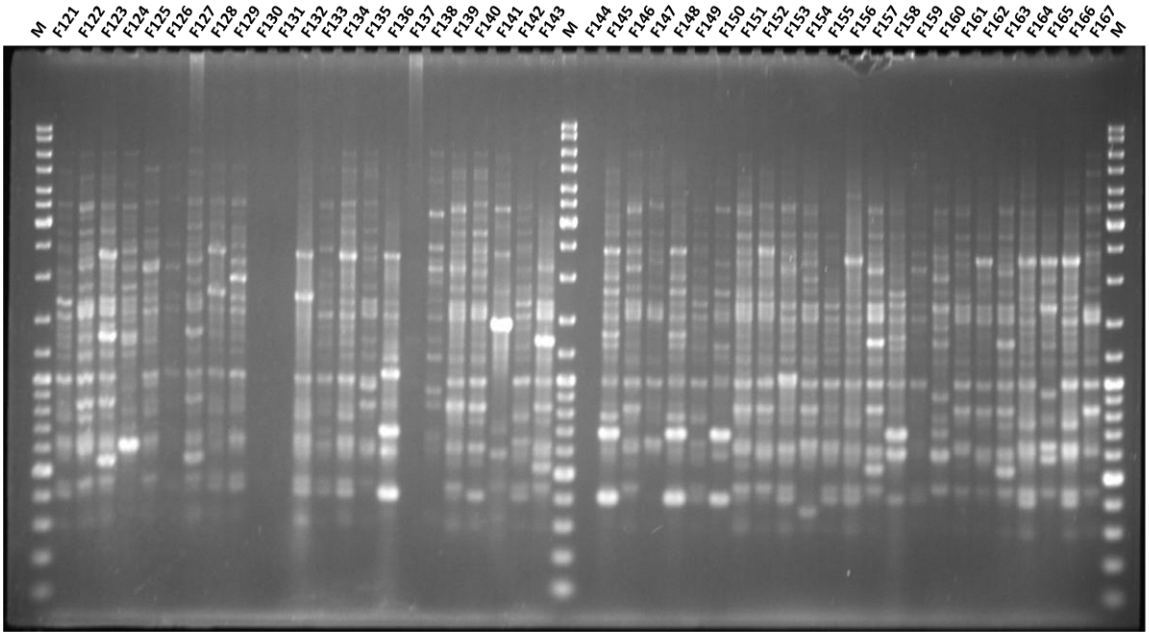


BOX-PCR 6:

- Top

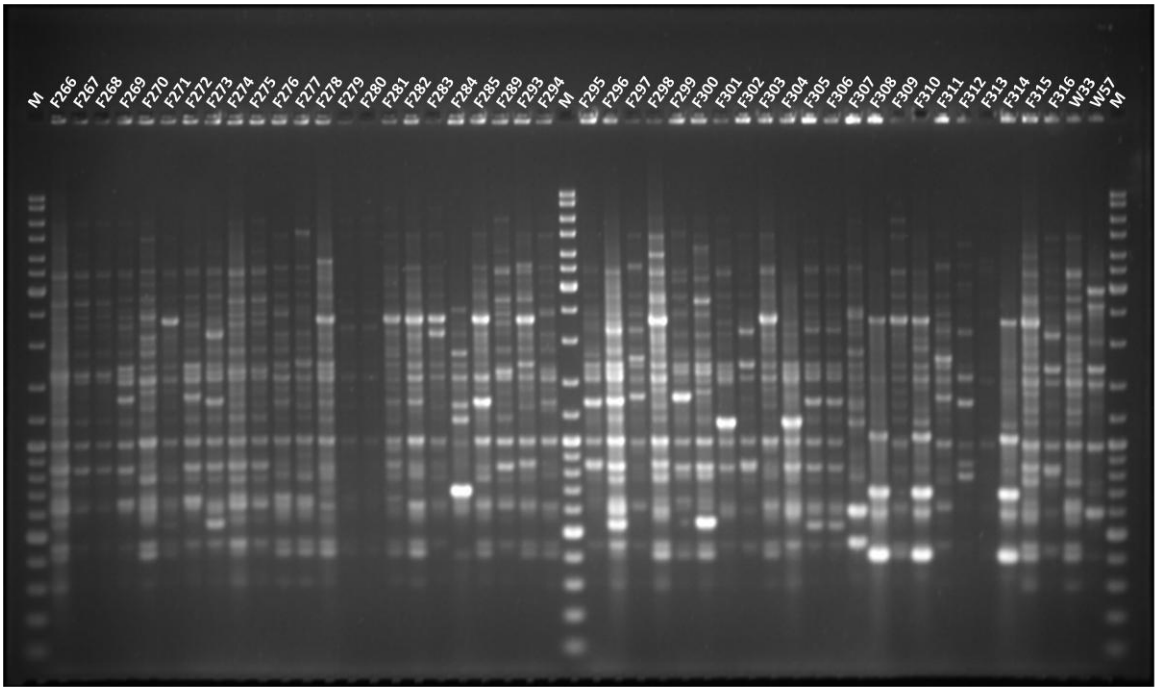


- Bottom

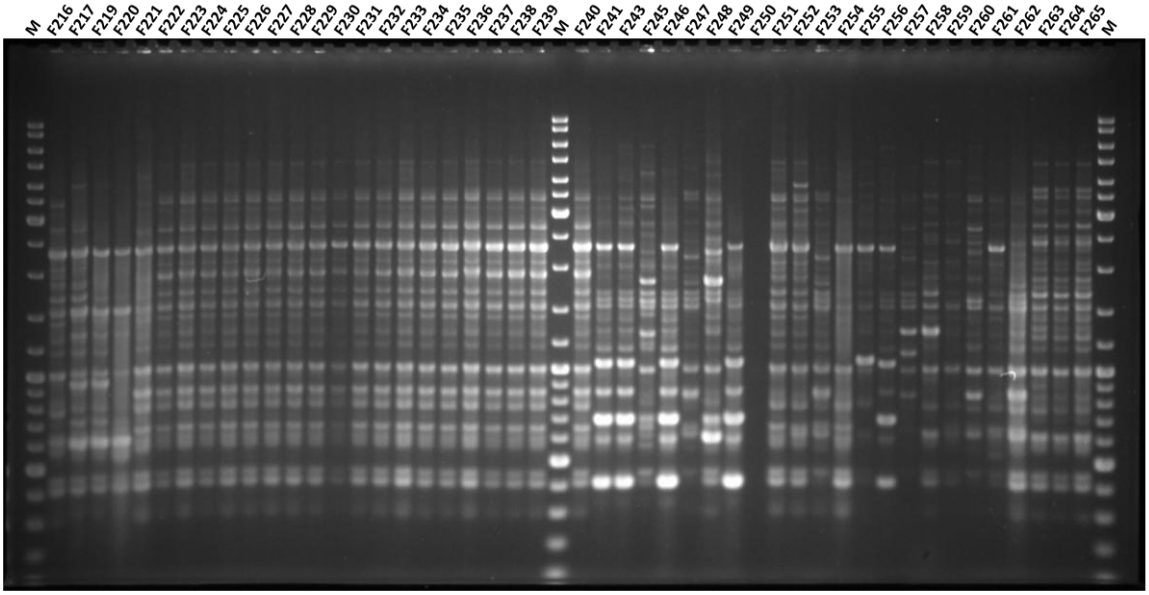


BOX-PCR 7:

- Top

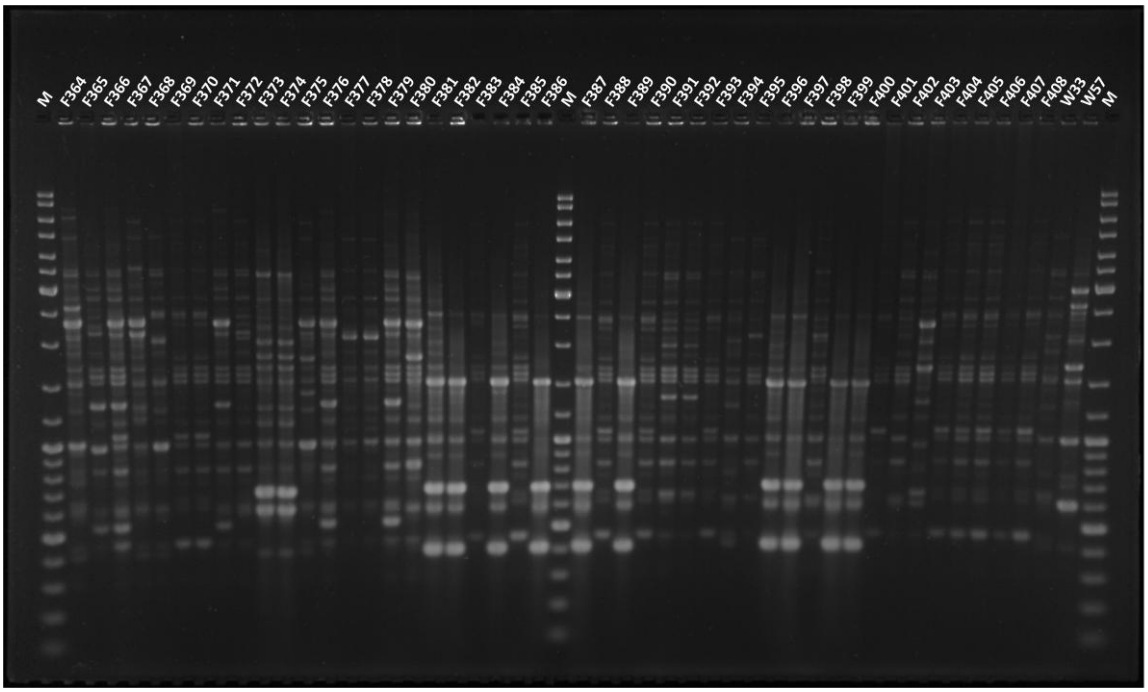


- Bottom

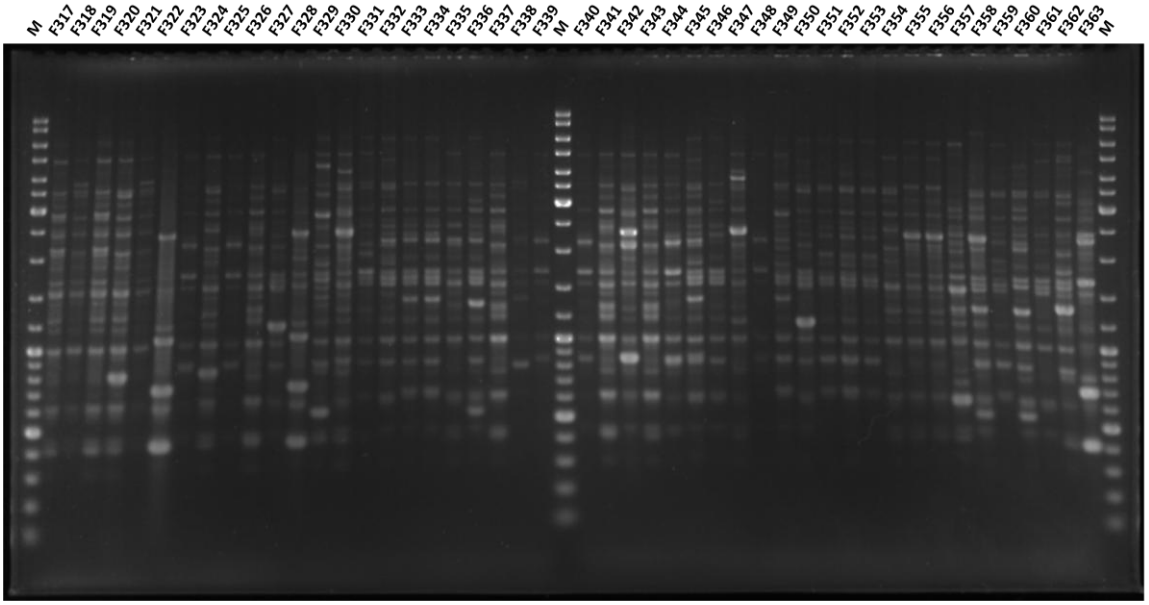


BOX-PCR 8:

- Top

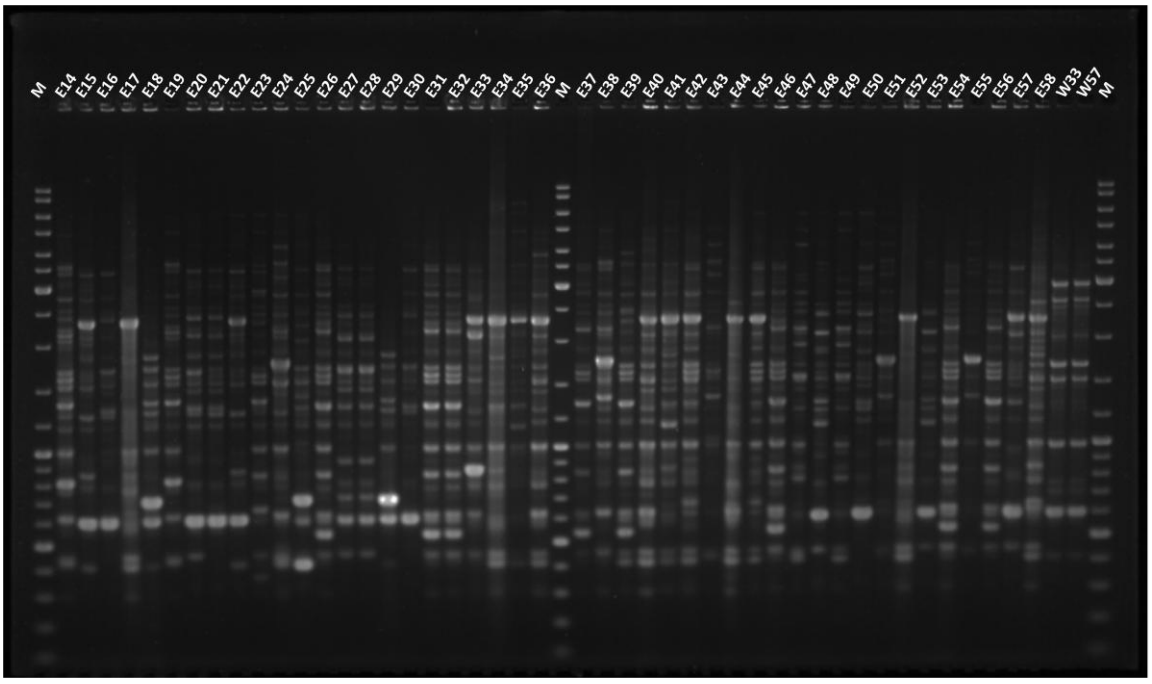


- Bottom

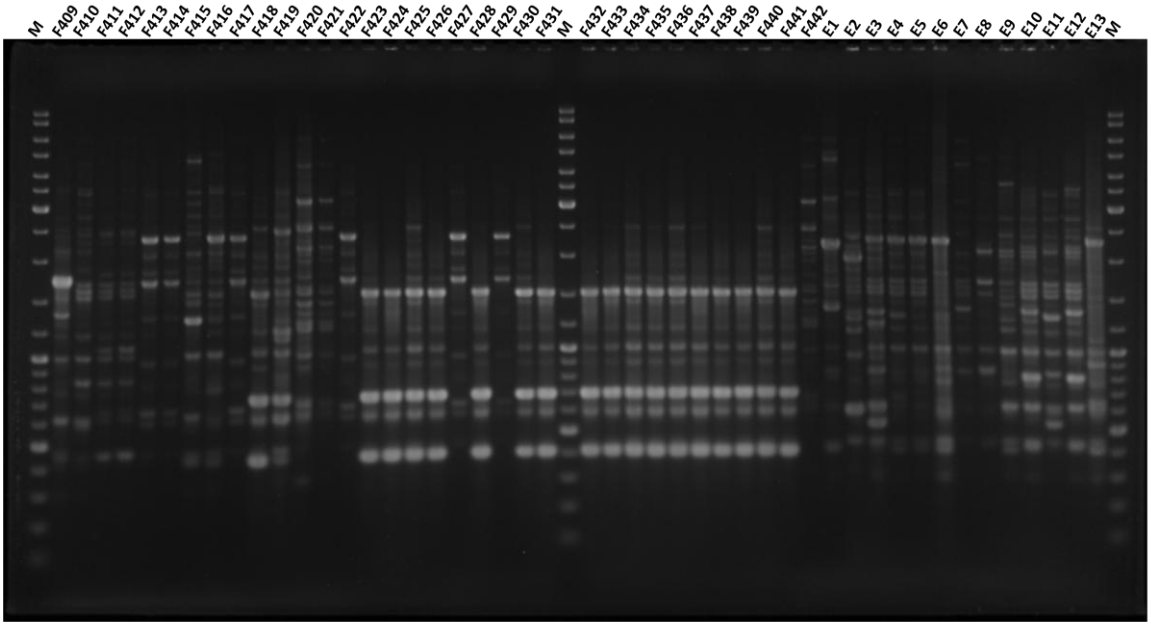


BOX-PCR 9:

- Top

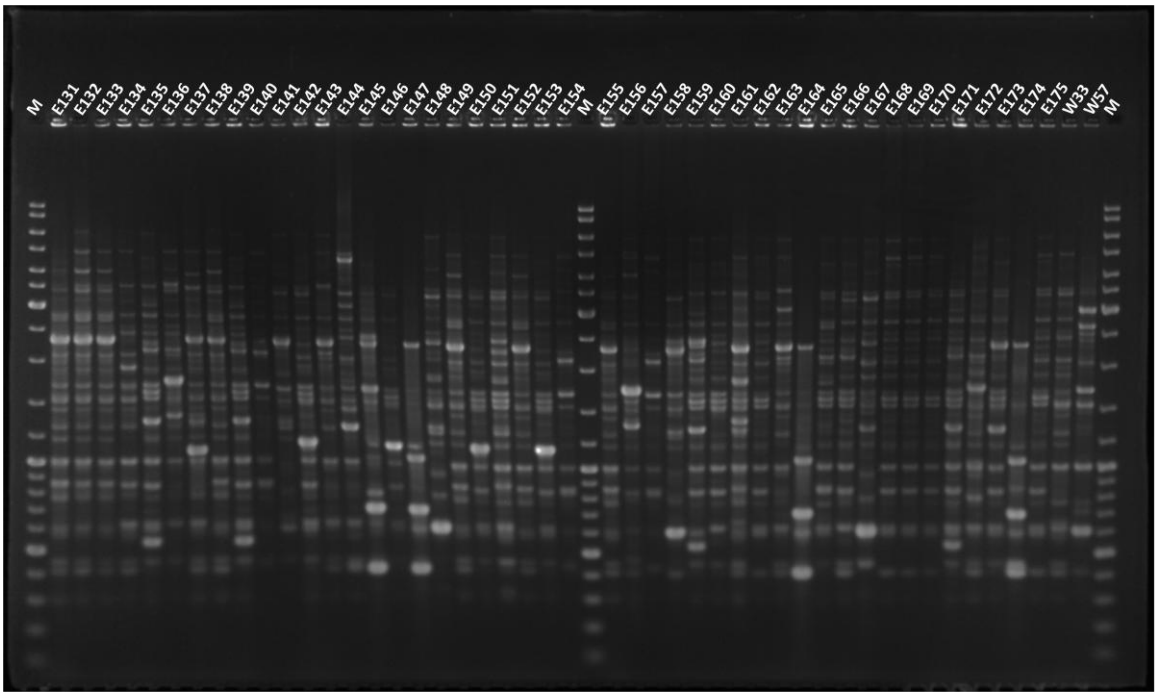


- Bottom

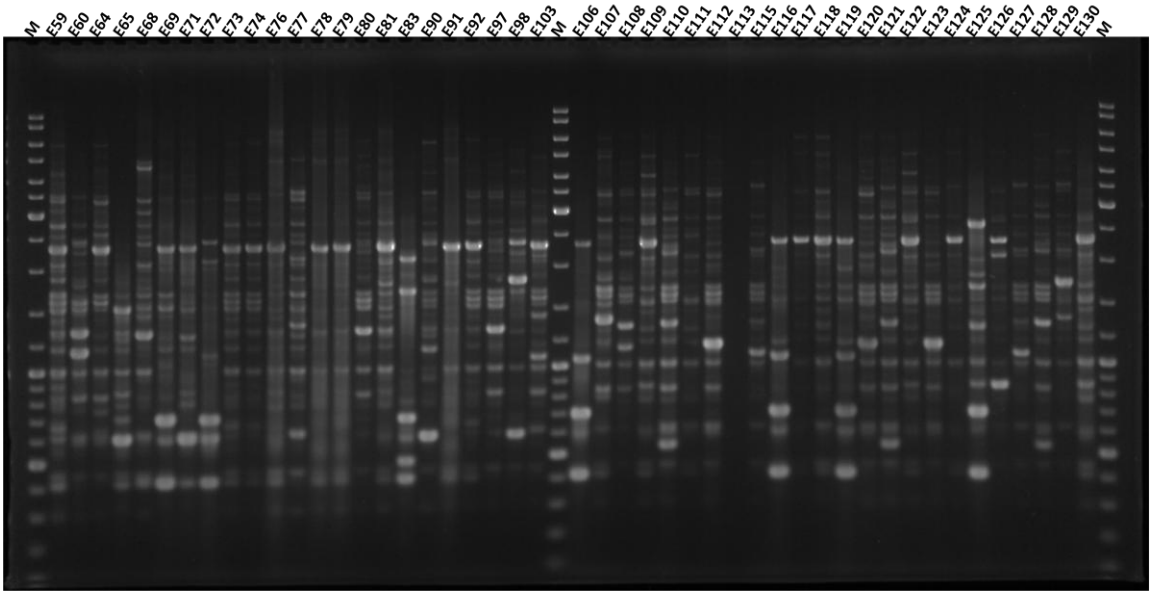


BOX-PCR 10:

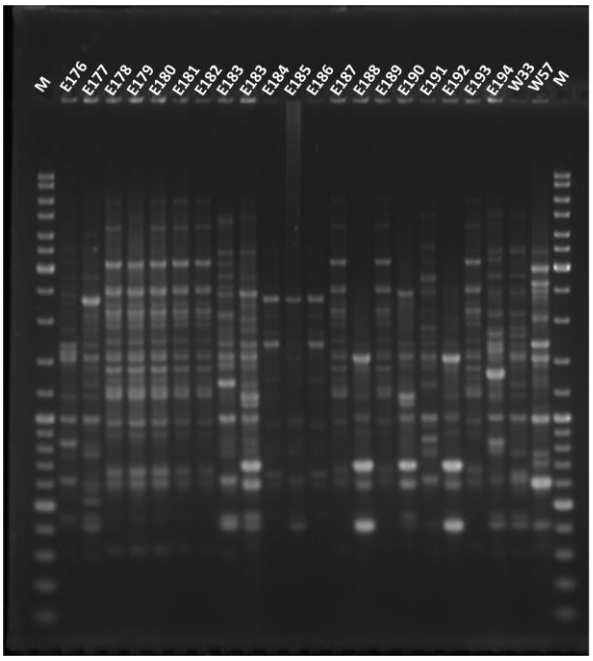
- Top



- Bottom



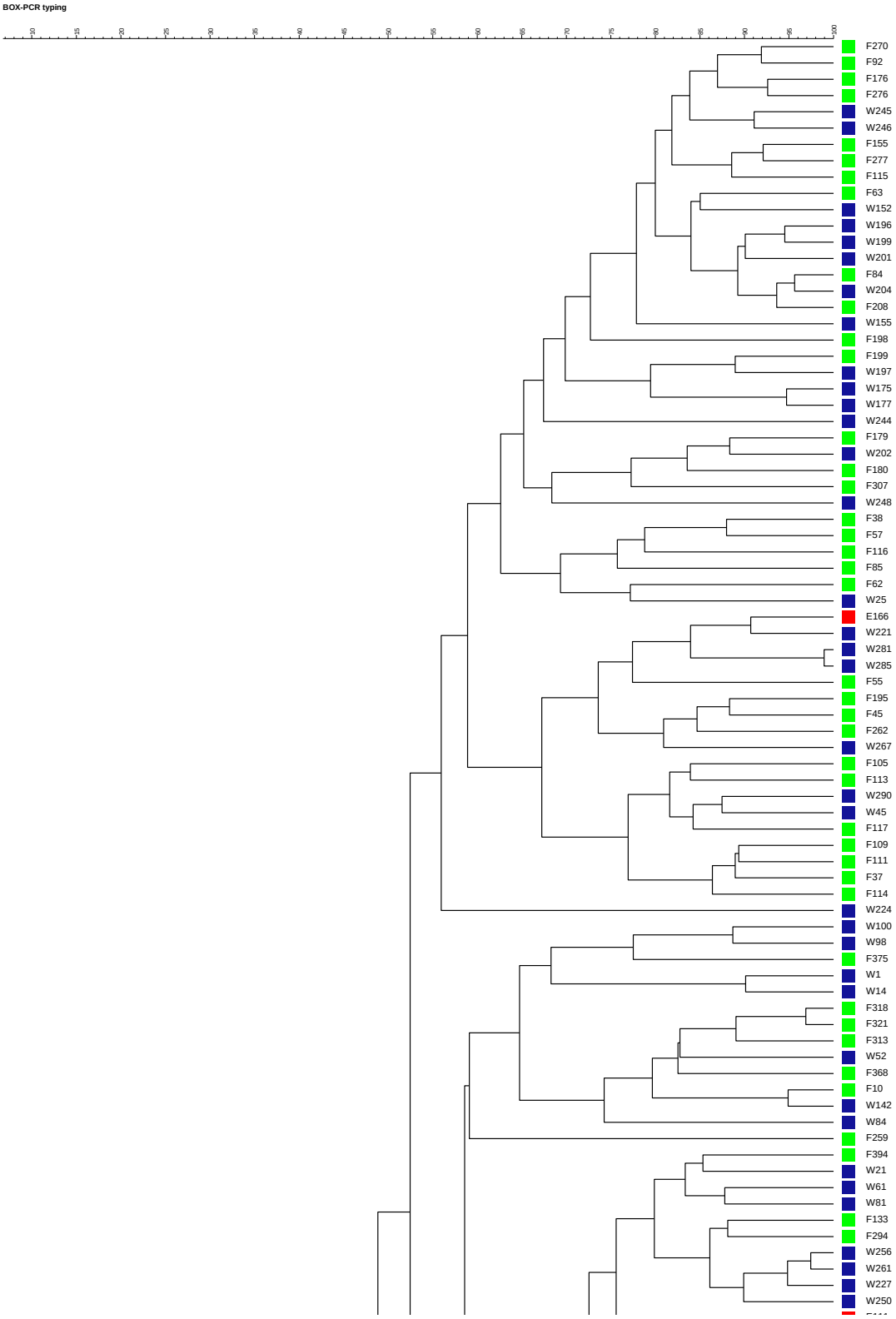
BOX-PCR 11:

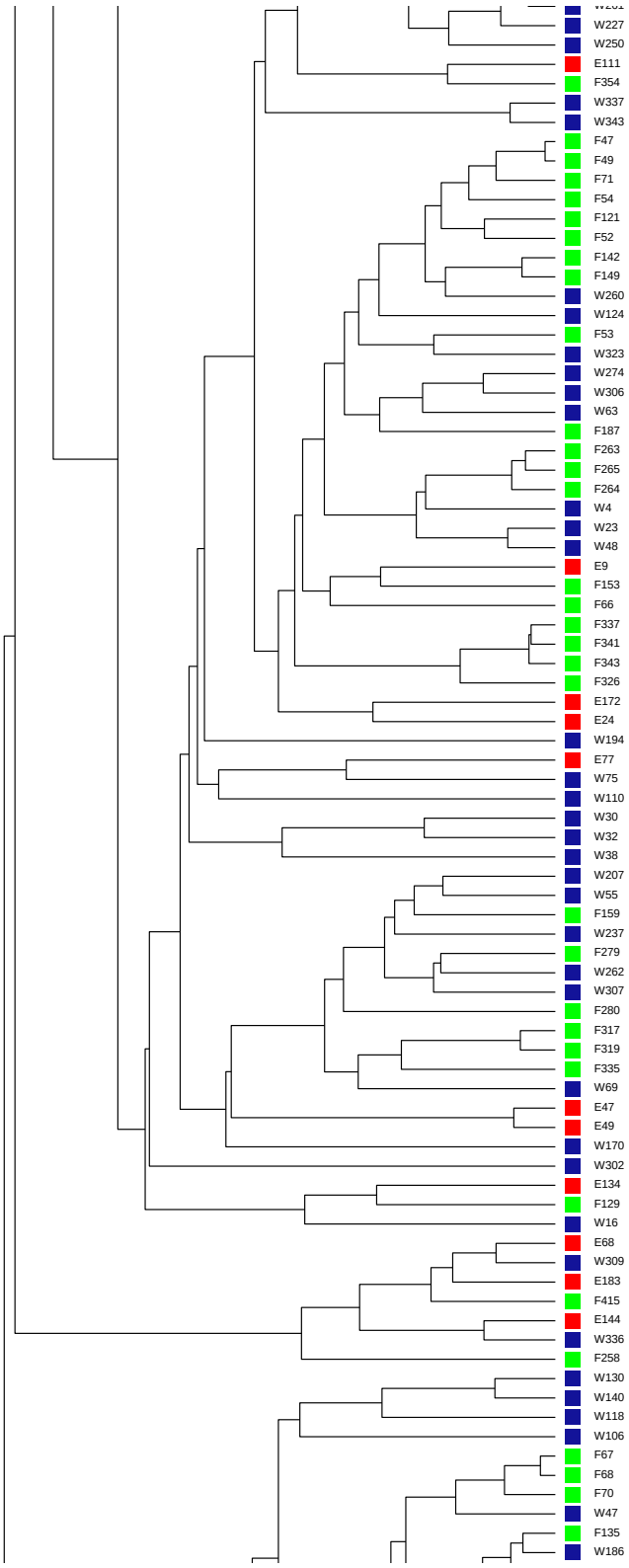


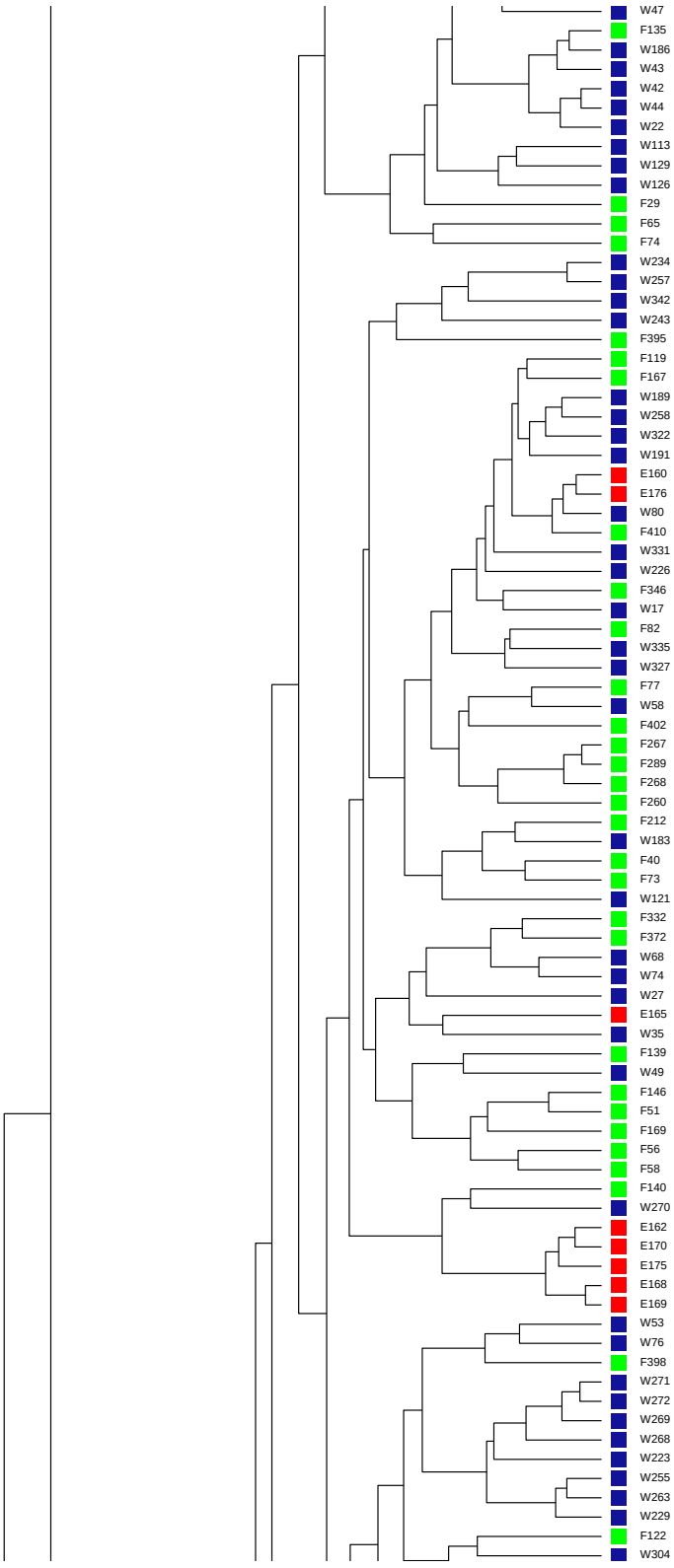
APPENDIX D – DENDROGRAMS

Apart from the designation above (appendix C), the three types of isolates collected are differentiated by a color, in the following dendrograms. Thus, the water isolates correspond to the blue squares, the effluent isolates correspond to the red squares, and the isolates derived from gull feces correspond to the green squares.

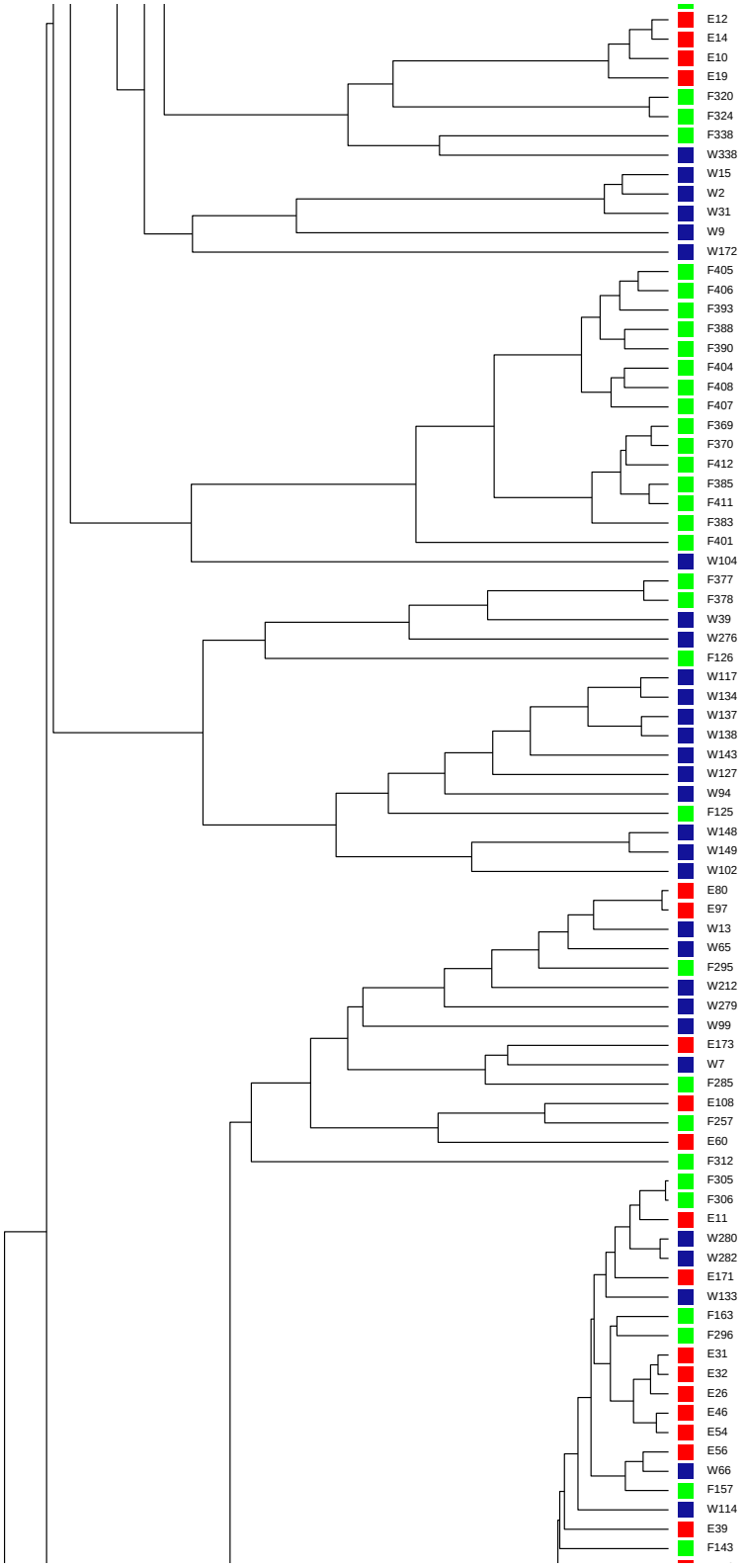
D.1 DENDROGRAM WITH ALL THE ISOLATES OF THE COLLECTION

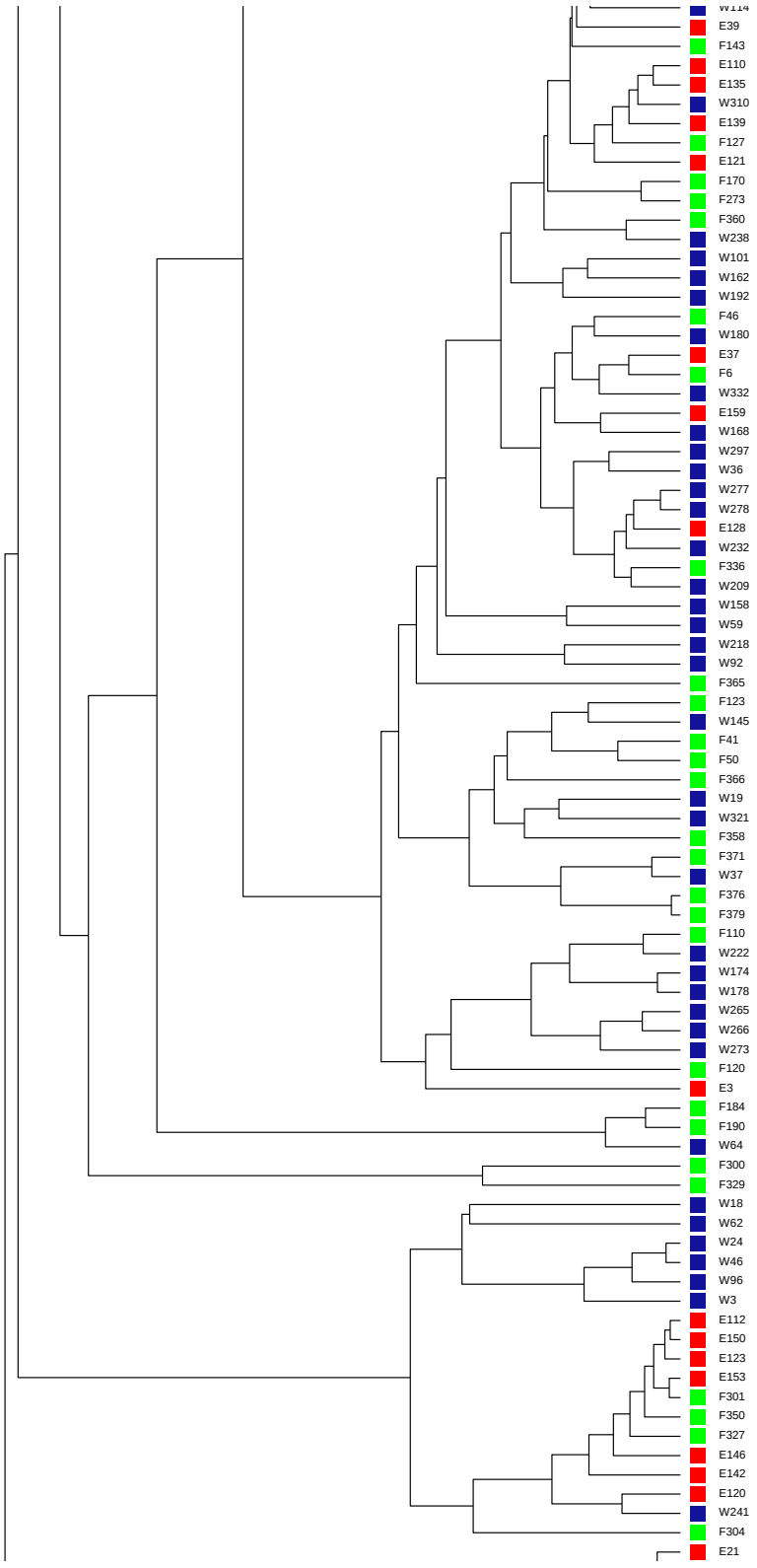

BOX-PCR typing
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
F270
F92
F176
F276
W245
W246
F155
F277
F115
F63
W152
W196
W199
W201
F84
W204
F208
W155
F198
F199
W197
W175
W177
W244
F179
W202
F180
F307
W248
F38
F57
F116
F85
F62
W25
E166
W221
W281
W285
F55
F195
F45
F262
W267
F105
F113
W290
W45
F117
F109
F111
F37
F114
W224
W100
W98
F375
W1
W14
F318
F321
F313
W52
F368
F10
W142
W84
F259
F394
W21
W61
W81
F133
F294
W256
W261
W227
W250

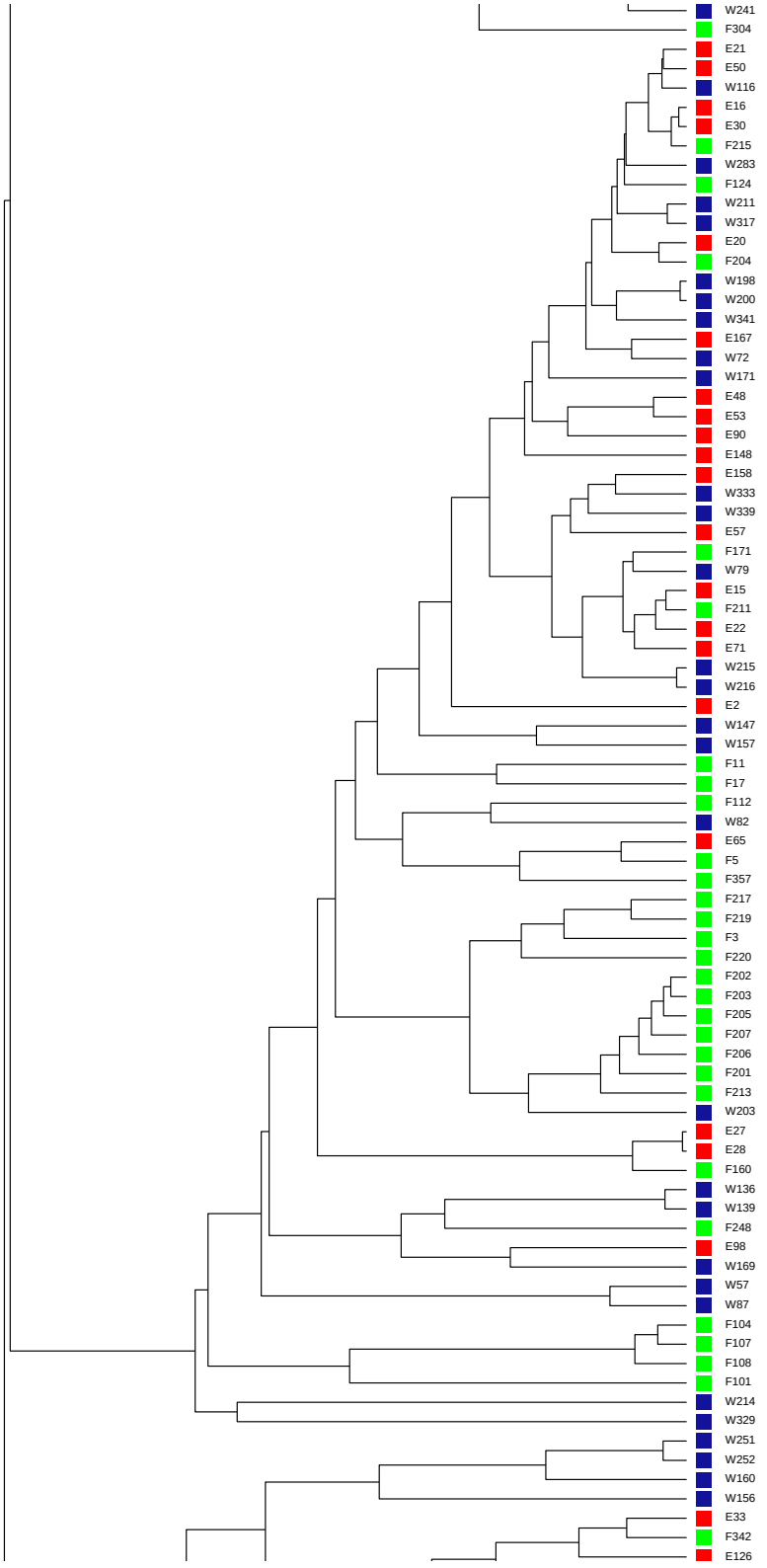

W227
W250
E111
F354
W337
W343
F47
F49
F71
F54
F121
F52
F142
F149
W260
W124
F53
W323
W274
W306
W63
F187
F263
F265
F264
W4
W23
W48
E9
F153
F66
F337
F341
F343
F326
E172
E24
W194
E77
W75
W110
W30
W32
W38
W207
W55
F159
W237
F279
W262
W307
F280
F317
F319
F335
W69
E47
E49
W170
W302
E134
F129
W16
E68
W309
E183
F415
E144
W336
F258
W130
W140
W118
W106
F67
F68
F70
W47
F135
W186

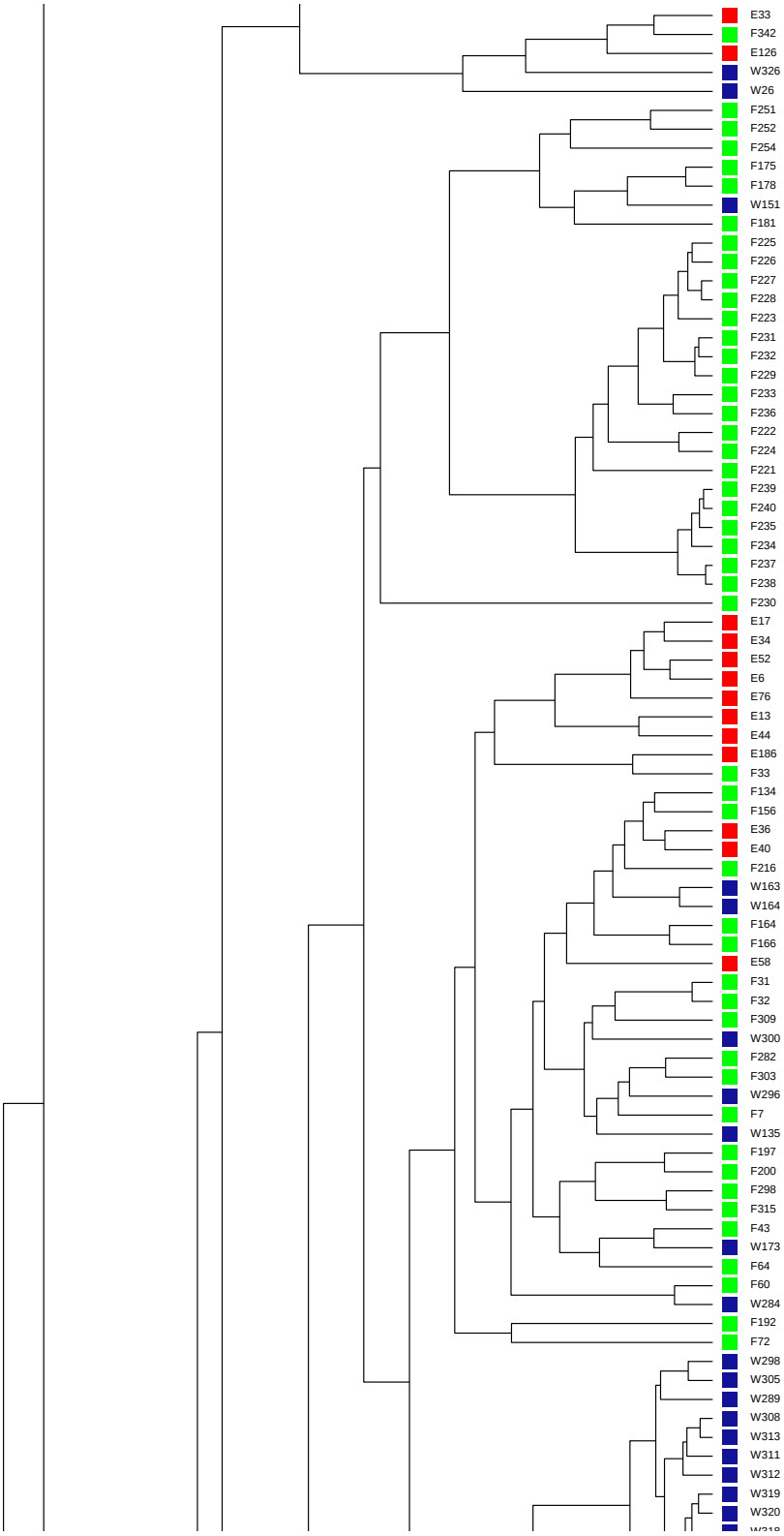

W47
F135
W186
W43
W42
W44
W22
W113
W129
W126
F29
F65
F74
W234
W257
W342
W243
F395
F119
F167
W189
W258
W322
W191
E160
E176
W80
F410
W331
W226
F346
W17
F82
W335
W327
F77
W58
F402
F267
F289
F268
F260
F212
W183
F40
F73
W121
F332
F372
W68
W74
W27
E165
W35
F139
W49
F146
F51
F169
F56
F58
F140
W270
E162
E170
E175
E168
E169
W53
W76
F398
W271
W272
W269
W268
W223
W255
W263
W229
F122
W304

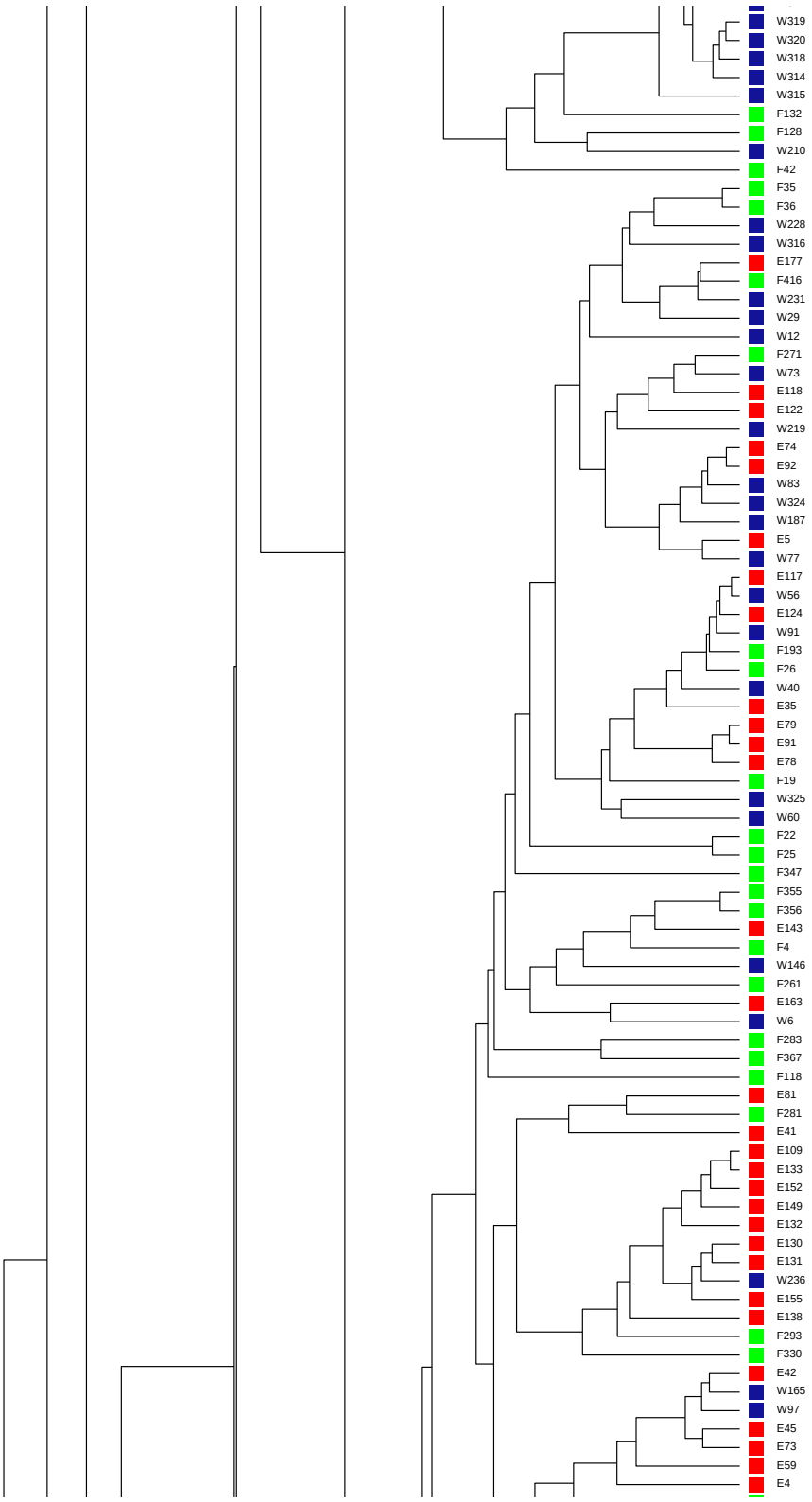

W229
F122
W304
W119
W188
W220
W240
W242
W112
F361
F1
W249
W159
F247
F253
F351
F353
F352
F359
E192
W108
W154
W247
W150
W193
F177
F86
F88
F81
F89
F90
W195
F185
F189
F183
F152
F266
F274
F87
F151
F161
F275
F147
W185
F196
W179
W33
F154
F172
W89
W90
W34
E151
W70
F191
W144
F269
F272
W78
F391
F392
F173
E23
F34
F349
F333
F334
F345
W11
W233
W235
E7
E195
F362
E107
F299
F15
F16
E12
E14

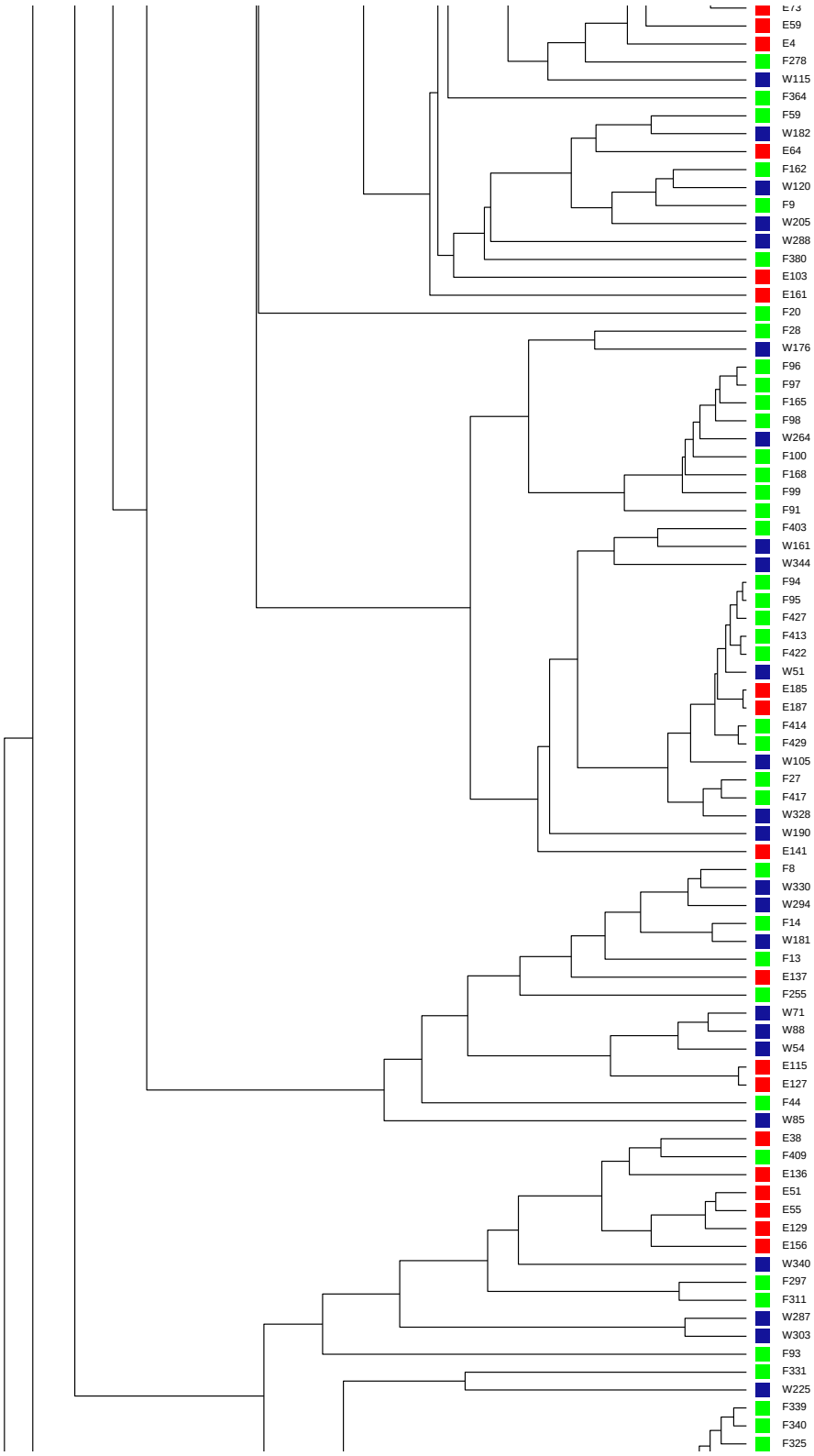

E12
E14
E10
E19
F320
F324
F338
W338
W15
W2
W31
W9
W172
F405
F406
F393
F388
F390
F404
F408
F407
F369
F370
F412
F385
F411
F383
F401
W104
F377
F378
W39
W276
F126
W117
W134
W137
W138
W143
W127
W94
F125
W148
W149
W102
E80
E97
W13
W65
F295
W212
W279
W99
E173
W7
F285
E108
F257
E60
F312
F305
F306
E11
W280
W282
E171
W133
F163
F296
E31
E32
E26
E46
E54
E56
W66
F157
W114
E39
F143

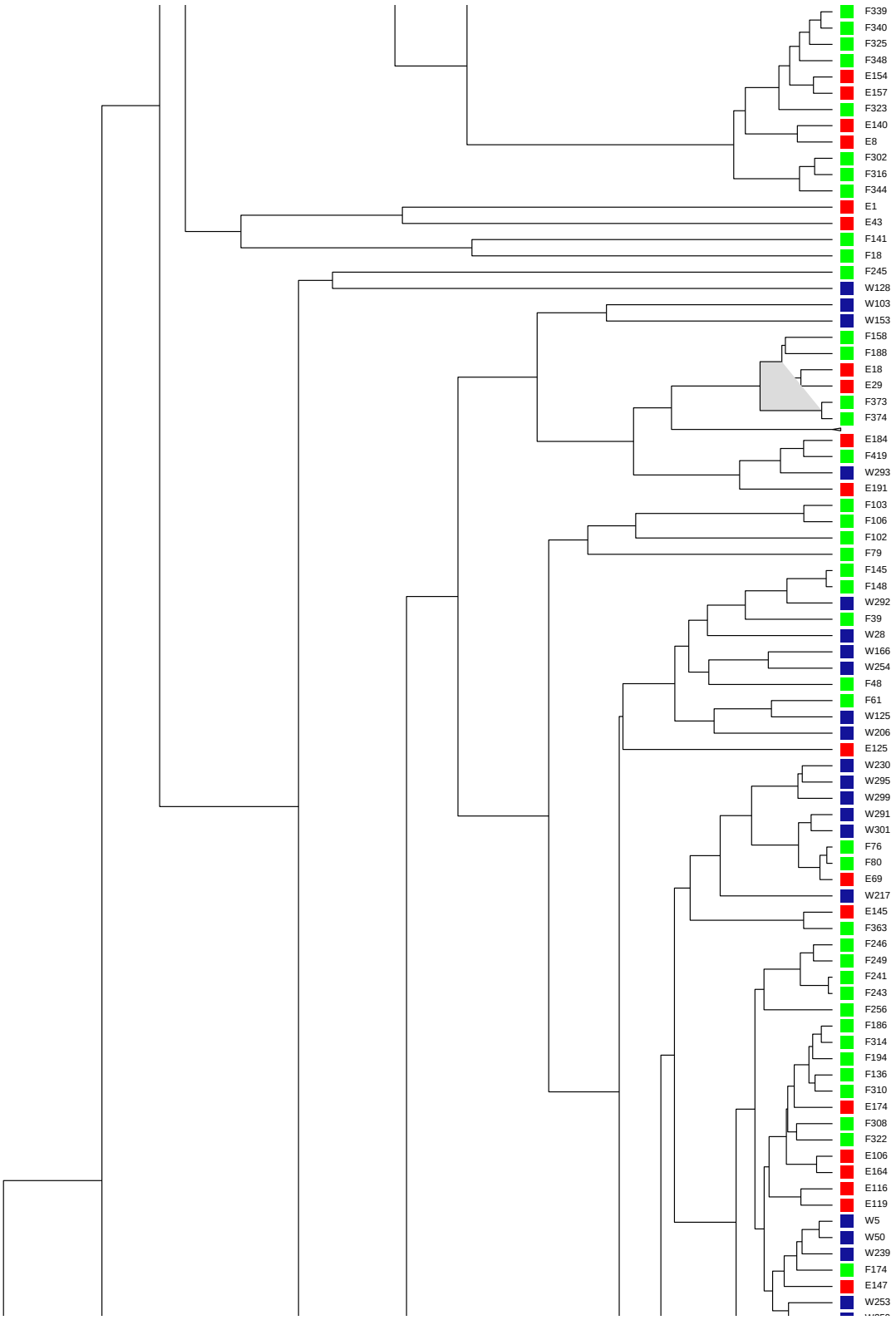

W114
E39
F143
E110
E135
W310
E139
F127
E121
F170
F273
F360
W238
W101
W162
W192
F46
W180
E37
F6
W332
E159
W168
W297
W36
W277
W278
E128
W232
F336
W209
W158
W59
W218
W92
F365
F123
W145
F41
F50
F366
W19
W321
F358
F371
W37
F376
F379
F110
W222
W174
W178
W265
W266
W273
F120
E3
F184
F190
W64
F300
F329
W18
W62
W24
W46
W96
W3
E112
E150
E123
E153
F301
F350
F327
E146
E142
E120
W241
F304
E21

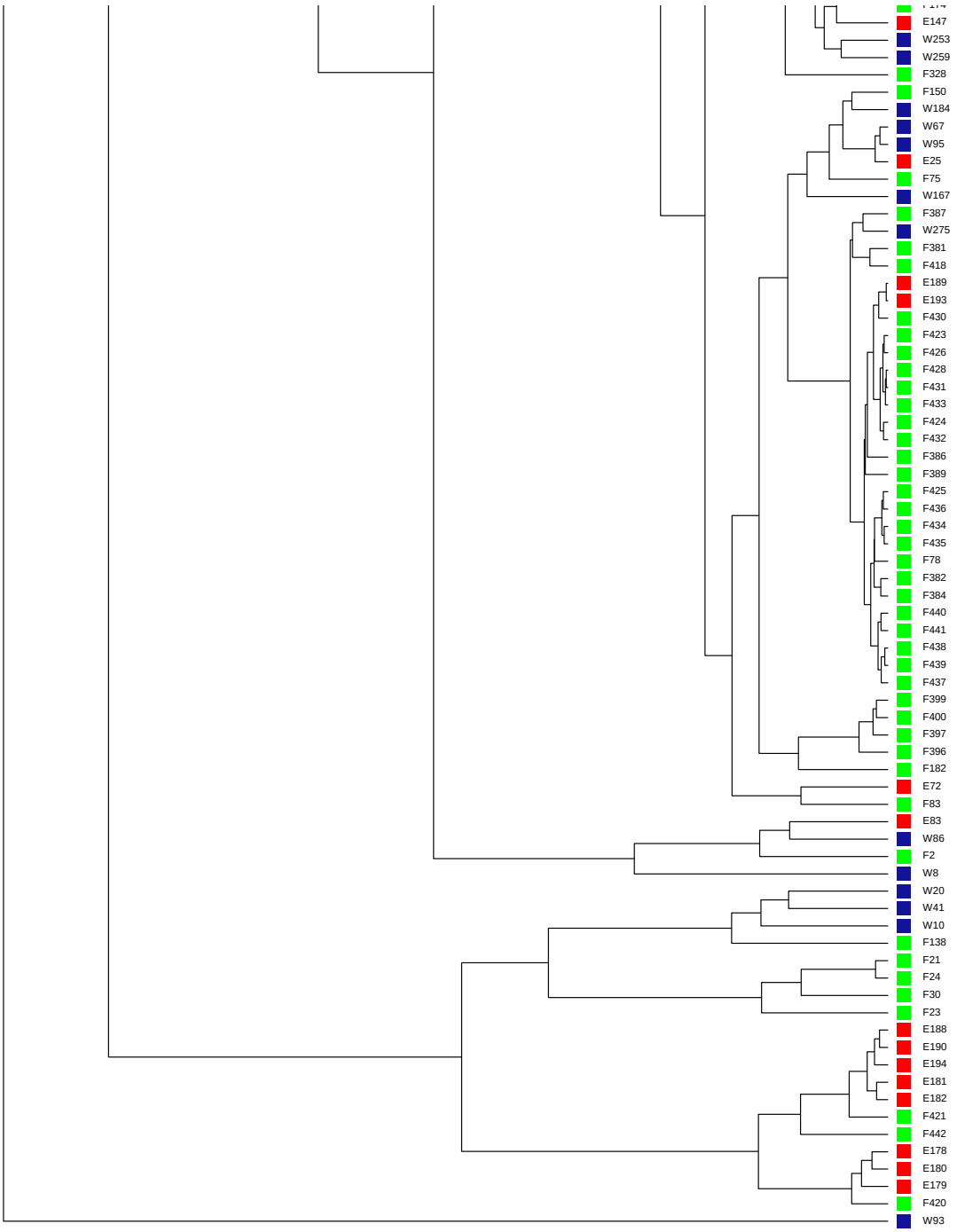

W241
F304
E21
E50
W116
E16
E30
F215
W283
F124
W211
W317
E20
F204
W198
W200
W341
E167
W72
W171
E48
E53
E90
E148
E158
W333
W339
E57
F171
W79
E15
F211
E22
E71
W215
W216
E2
W147
W157
F11
F17
F112
W82
E65
F5
F357
F217
F219
F3
F220
F202
F203
F205
F207
F206
F201
F213
W203
E27
E28
F160
W136
W139
F248
E98
W169
W57
W87
F104
F107
F108
F101
W214
W329
W251
W252
W160
W156
E33
F342
E126


E33
F342
E126
W326
W26
F251
F252
F254
F175
F178
W151
F181
F225
F226
F227
F228
F223
F231
F232
F229
F233
F236
F222
F224
F221
F239
F240
F235
F234
F237
F238
F230
E17
E34
E52
E6
E76
E13
E44
E186
F33
F134
F156
E36
E40
F216
W163
W164
F164
F166
E58
F31
F32
F309
W300
F282
F303
W296
F7
W135
F197
F200
F298
F315
F43
W173
F64
F60
W284
F192
F72
W298
W305
W289
W308
W313
W311
W312
W319
W320


W319
W320
W318
W314
W315
F132
F128
W210
F42
F35
F36
W228
W316
E177
F416
W231
W29
W12
F271
W73
E118
E122
W219
E74
E92
W83
W324
W187
E5
W77
E117
W56
E124
W91
F193
F26
W40
E35
E79
E91
E78
F19
W325
W60
F22
F25
F347
F355
F356
E143
F4
W146
F261
E163
W6
F283
F367
F118
E81
F281
E41
E109
E133
E152
E149
E132
E130
E131
W236
E155
E138
F293
F330
E42
W165
W97
E45
E73
E59
E4

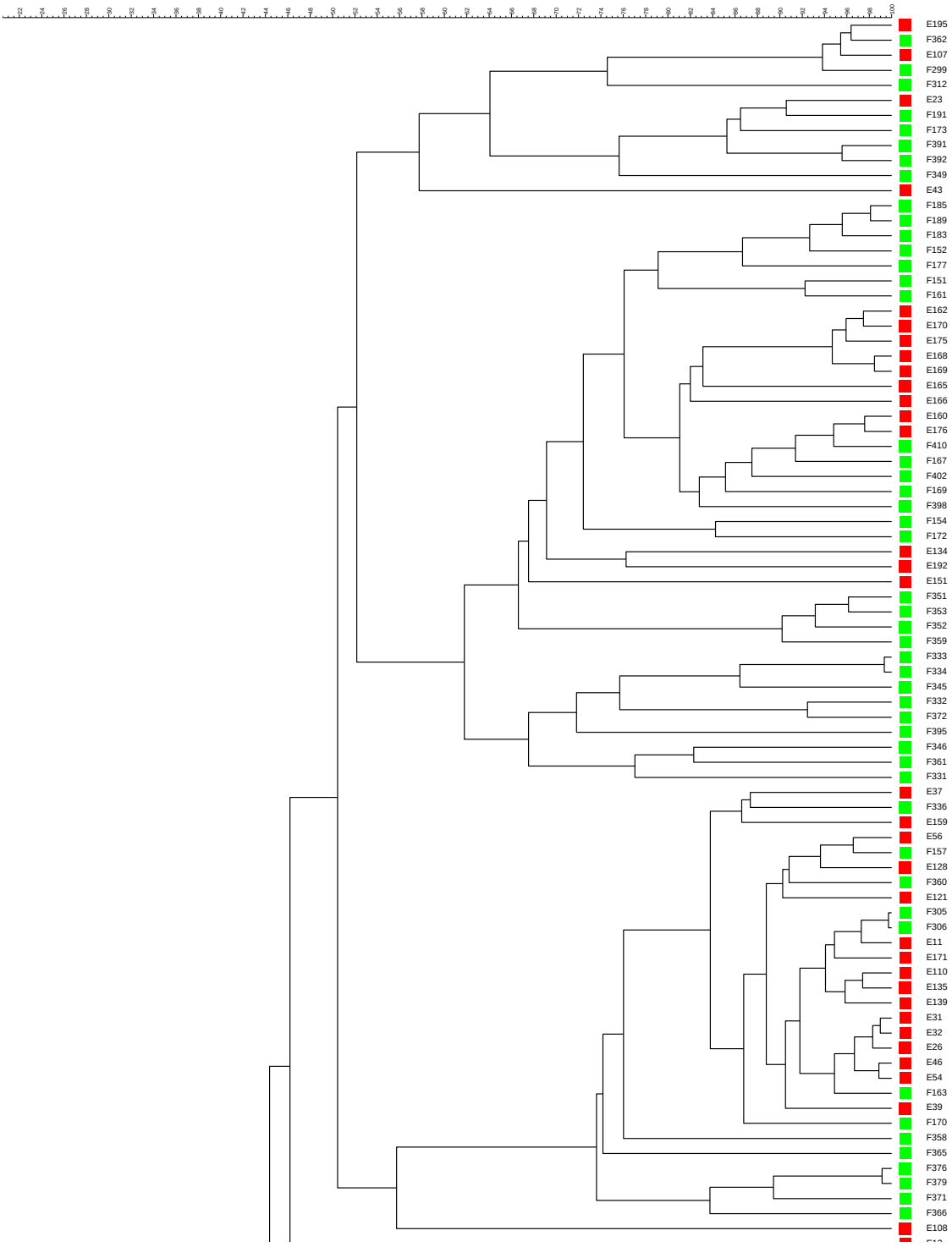

E73
E59
E4
F278
W115
F364
F59
W182
E64
F162
W120
F9
W205
W288
F380
E103
E161
F20
F28
W176
F96
F97
F165
F98
W264
F100
F168
F99
F91
F403
W161
W344
F94
F95
F427
F413
F422
W51
E185
E187
F414
F429
W105
F27
F417
W328
W190
E141
F8
W330
W294
F14
W181
F13
E137
F255
W71
W88
W54
E115
E127
F44
W85
E38
F409
E136
E51
E55
E129
E156
W340
F297
F311
W287
W303
F93
F331
W225
F339
F340
F325

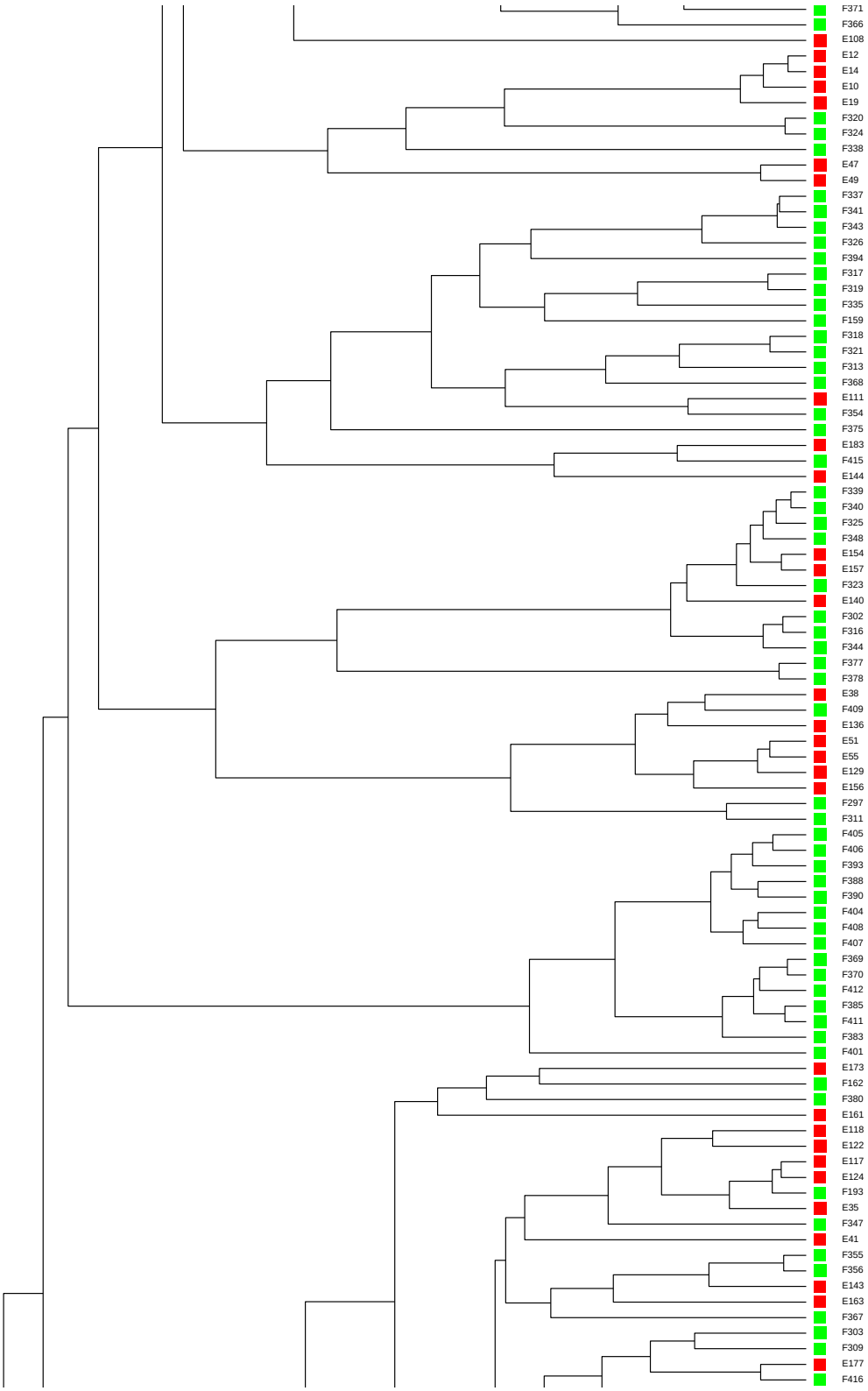

F339
F340
F325
F348
E154
E157
F323
E140
E8
F302
F316
F344
E1
E43
F141
F18
F245
W128
W103
W153
F158
F188
E18
E29
F373
F374
E184
F419
W293
E191
F103
F106
F102
F79
F145
F148
W292
F39
W28
W166
W254
F48
F61
W125
W206
E125
W230
W295
W299
W291
W301
F76
F80
E69
W217
E145
F363
F246
F249
F241
F243
F256
F186
F314
F194
F136
F310
E174
F308
F322
E106
E164
E116
E119
W5
W50
W239
F174
E147
W253

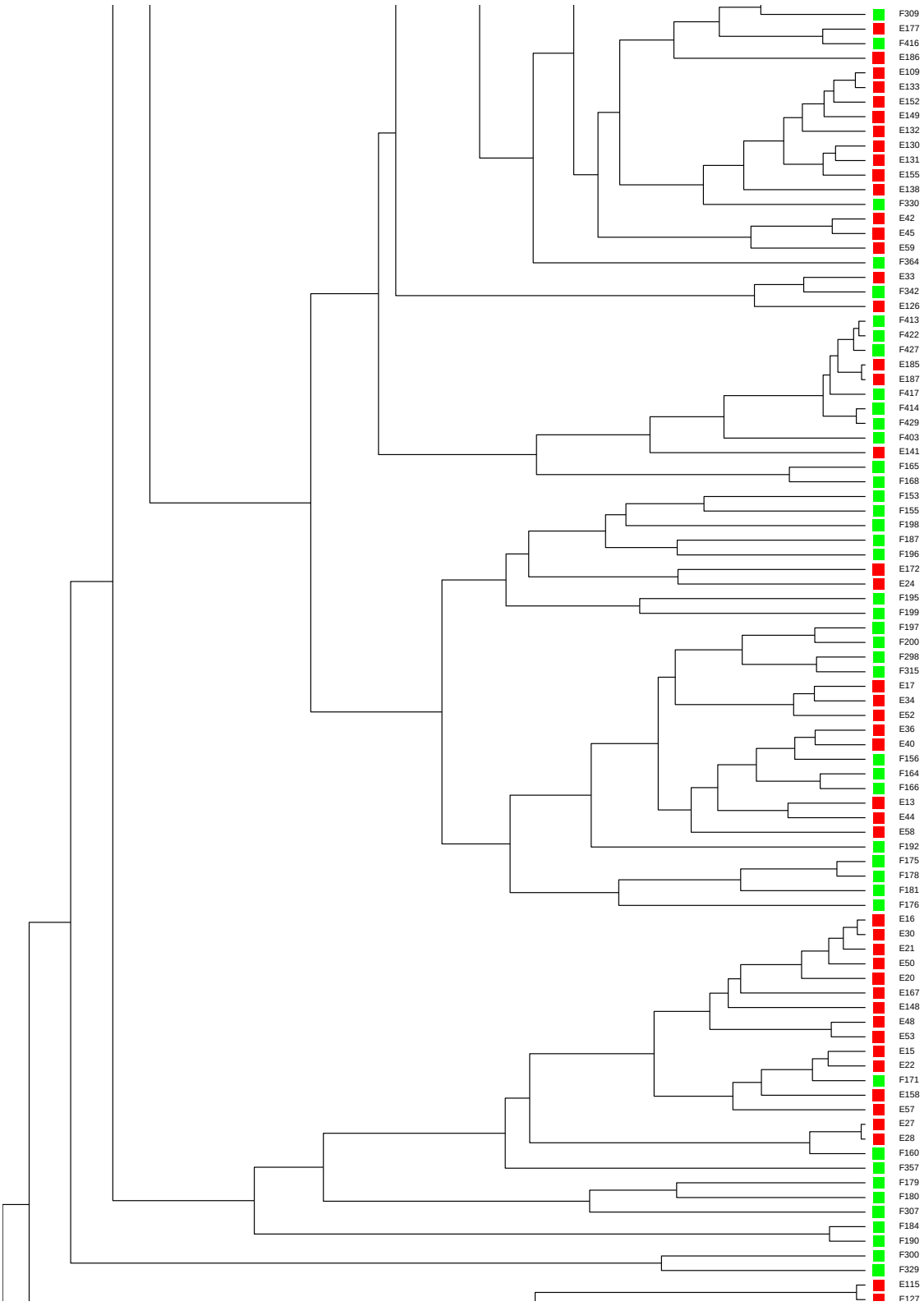

E147
W253
W259
F328
F150
W184
W67
W95
E25
F75
W167
F387
W275
F381
F418
E189
E193
F430
F423
F426
F428
F431
F433
F424
F432
F386
F389
F425
F436
F434
F435
F78
F382
F384
F440
F441
F438
F439
F437
F399
F400
F397
F396
F182
E72
F83
E83
W86
F2
W8
W20
W41
W10
F138
F21
F24
F30
F23
E188
E190
E194
E181
E182
F421
F442
E178
E180
E179
F420
W93

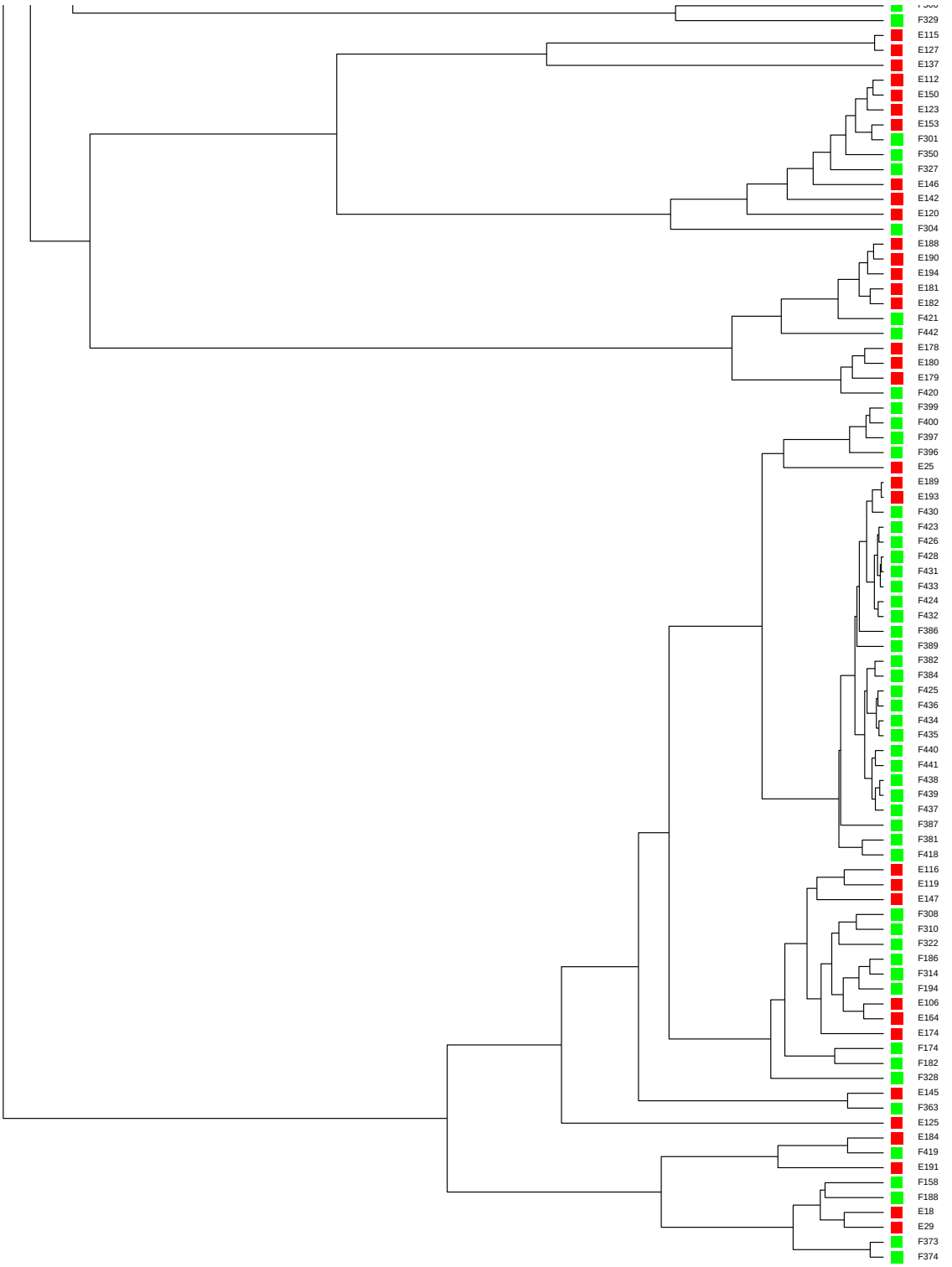
D.2 DENDROGRAM WITH FECES AND EFFLUENT ISOLATES

Pearson correlation (Opt:1.00%) [0.0%-100.0%]
BOX-PCR typing




F371
F366
E108
E12
E14
E10
E19
F320
F324
F338
E47
E49
F337
F341
F343
F326
F394
F317
F319
F335
F159
F318
F321
F313
F368
E111
F354
F375
E183
F415
E144
F339
F340
F325
F348
E154
E157
F323
E140
F302
F316
F344
F377
F378
E38
F409
E136
E51
E55
E129
E156
F297
F311
F405
F406
F393
F388
F390
F404
F408
F407
F369
F370
F412
F385
F411
F383
F401
E173
F162
F380
E161
E118
E122
E117
E124
F193
E35
F347
E41
F355
F356
E143
E163
F367
F303
F309
E177
F416


F309
E177
F416
E186
E109
E133
E152
E149
E132
E130
E131
E155
E138
F330
E42
E45
E59
F364
E33
F342
E126
F413
F422
F427
E185
E187
F417
F414
F429
F403
E141
F165
F168
F153
F155
F198
F187
F196
E172
E24
F195
F199
F197
F200
F298
F315
E17
E34
E52
E36
E40
F156
F164
F166
E13
E44
E58
F192
F175
F178
F181
F176
E16
E30
E21
E50
E20
E167
E148
E48
E53
E15
E22
F171
E158
E57
E27
E28
F160
F357
F179
F180
F307
F184
F190
F300
F329
E115
E127


F329
E115
E127
E137
E112
E150
E123
E153
F301
F350
F327
E146
E142
E120
F304
E188
E190
E194
E181
E182
F421
F442
E178
E180
E179
F420
F399
F400
F397
F396
E25
E189
E193
F430
F423
F426
F428
F431
F433
F424
F432
F386
F389
F382
F384
F425
F436
F434
F435
F440
F441
F438
F439
F437
F387
F381
F418
E116
E119
E147
F308
F310
F322
F186
F314
F194
E106
E164
E174
F174
F182
F328
E145
F363
E125
E184
F419
E191
F158
F188
E18
E29
F373
F374